Geophysical Journal International

Geophys. J. Int. (2010) **183,** 1014–1030 doi: 10.1111/j.1365-246X.2010.04774.x

Probabilistic prediction of rupture length, slip and seismic ground motions for an ongoing rupture: implications for early warning for large earthquakes

Maren Böse and Thomas H. Heaton

Seismological Laboratory, California Institute of Technology (Caltech), 1200 *E. California Blvd., Pasadena, CA* 91125, *USA. E-mail: mboese@caltech.edu*

Accepted 2010 August 12. Received 2010 August 11; in original form 2010 February 12

SUMMARY

Earthquake Early Warning (EEW) predicts future ground shaking based on presently available data. Long ruptures present the best opportunities for EEW since many heavily shaken areas are distant from the earthquake epicentre and may receive long warning times. Predicting the shaking from large earthquakes, however, requires some estimate of the likelihood of the future evolution of an ongoing rupture. An EEW system that anticipates future rupture using the present magnitude (or rupture length) together with the Gutenberg-Richter frequency-size statistics will likely never predict a large earthquake, because of the rare occurrence of 'extreme events'. However, it seems reasonable to assume that large slip amplitudes increase the probability for evolving into a large earthquake. To investigate the relationship between the slip and the eventual size of an ongoing rupture, we simulate suites of 1-D rupture series from stochastic models of spatially heterogeneous slip. We find that while large slip amplitudes increase the probability for the continuation of a rupture and the possible evolution into a 'Big One', the recognition that rupture is occurring on a spatially smooth fault has an even stronger effect. We conclude that an EEW system for large earthquakes needs some mechanism for the rapid recognition of the causative fault (e.g., from real-time GPS measurements) and consideration of its 'smoothness'. An EEW system for large earthquakes on smooth faults, such as the San Andreas Fault, could be implemented in two ways: the system could issue a warning, whenever slip on the fault exceeds a few metres, because the probability for a large earthquake is high and strong shaking is expected to occur in large areas around the fault. A more sophisticated EEW system could use the present slip on the fault to estimate the future slip evolution and final rupture dimensions, and (using this information) could provide probabilistic predictions of seismic ground motions along the evolving rupture. The decision on whether an EEW system should be realized in the first or in the second way (or in a combination of both) is user-specific.

Key words: Probabilistic forecasting; Earthquake dynamics; Earthquake ground motions; Seismicity and tectonics; Statistical seismology; Early warning.

INTRODUCTION

Earthquake Early Warning (EEW) requires fast and robust predictions of earthquake source and ground motion parameters shortly after the initiation of an earthquake. If given in a timely manner, warnings can be used to trigger and execute automatic measures to reduce expected damage at distant sites (e.g. Goltz 2002; Allen *et al.* 2009). Several algorithms for EEW have been proposed and tested during recent years (e.g. Nakamura 1988; Allen & Kanamori 2003; Kanamori 2005; Böse 2006; Böse *et al.* 2007; Cua & Heaton 2007; Wu *et al.* 2007; Hoshiba *et al.* 2008; Böse *et al.* 2009a,b). The majority of these approaches use the amplitude and frequency content of seismic waves at one or more seismic sensors in the epicentral area for a rapid assessment of earthquake source and ground motion parameters. With a few exceptions (e.g. Yamada *et al.* 2007; Böse *et al.* 2008; Yamada & Heaton 2008), the majority of the current EEW approaches consider the earthquake as a point source, that is, neglect rupture finiteness.

Large magnitude earthquakes ($M > 7.0$), with rupture lengths of up to hundreds of kilometres, cause damaging ground shaking in much larger areas than moderate-to-strong events. Despite their rare occurrence, many more people and users could benefit from an EEW system during large earthquakes (Heaton 1985; Allen 2006). Due to the low propagation speed of seismic ruptures (∼80 per cent of the seismic *S*-wave speed), warning times to heavily shaken areas may exceed more than 1 min, while users who are strongly shaken

© 2010 The Authors
Geophysical Journal International © 2010 RAS

in smaller earthquakes are likely to be near the epicentre and will generally receive warnings of a few seconds only.

A key to EEW for large magnitude earthquakes is the rapid prediction of the potential final dimensions of the ongoing rupture. This is because the level and distribution of seismic ground motions, such as peak ground velocity (PGV) or seismic intensity, are usually predicted from the earthquake magnitude and the rupture-to-site distance, which both can be estimated from the expected rupture dimensions. If the rupture is still propagating, these predictions can only be given in a probabilistic manner. Appropriate decision making is necessarily based on probabilistic estimates of future shaking.

The rupture of a large (strike-slip) earthquake propagates mainly in the horizontal direction along a fault, that is, rapid prediction of the rupture length is especially important for EEW. An empirical relationship between the subsurface rupture length L(km) and the (moment) magnitude M for strike-slip earthquakes with M 4.8 to M 8.1, for example, has been proposed by Wells & Coppersmith (1994) with

$$M = 4.33 + 1.49 \log(L) \tag{1}$$

[see Working Group on California Earthquake Probabilities (1999) for discussions on this relation].

A simple way to predict L (without any further information on the rupture) is by the usage of frequency-size statistical distributions of earthquakes, such as produced by the Gutenberg-and-Richter model (G-R; Gutenberg & Richter 1944) or by the Characteristic Earthquake model (CE; Schwartz & Coppersmith 1984).

The G-R model describes the frequency-size distribution of earthquakes by the power-law

$$\log N(M) = a - bM, \tag{2}$$

where $N(M)$ is the number of earthquakes exceeding magnitude M, and a and b are constants (Gutenberg & Richter 1944). The b-value in eq. (2) is usually observed to be about 1.0, although there are regional variations (e.g. Wiemer & Wyss 2002; Schorlemmer *et al.* 2004); b also depends on the particular definition of magnitude chosen (e.g. $M_{\rm l}$, $M_{\rm S}$ and $M_{\rm w}$).

The probability of occurrence of an earthquake exceeding a certain magnitude M in the range $[M_{\rm min}, M_{\rm max}]$ can be described by the probability density function (PDF)

$$\begin{aligned}\mathrm{prob}_{G-R}(M) &= \frac{N(M)}{\int_{M_p}^{M_{\rm max}} N(M)\,\mathrm{d}M} = \frac{\beta \mathrm{e}^{-\beta M}}{\mathrm{e}^{-\beta M_{\rm min}} - \mathrm{e}^{-\beta M_{\rm max}}} \\ &= \frac{\beta \mathrm{e}^{-\beta(M-M_{\rm min})}}{1-\mathrm{e}^{-\beta(M_{\rm max}-M_{\rm min})}},\end{aligned} \tag{3}$$

with $\beta = b \ln[10]$ and $\int_{M_{\rm min}}^{M_{\rm max}} \mathrm{prob}_{G\text{-}R}(M)\mathrm{d}M = 1$ (Fig. 1a). The probability of exceedance (POE) of a present magnitude M_p is

$$p_{G-R}(M|M_p) = \frac{\mathrm{prob}_{G\text{-}R}(M)}{\mathrm{prob}_{G\text{-}R}(M_p)} = \mathrm{e}^{-\beta(M-M_p)} = 10^{-b(M-M_p)}. \tag{4}$$

If, for example, we knew that the present magnitude of an ongoing earthquake was M_p 5.5 (corresponding to a present rupture length of $L_p \approx 6$ km) and $b = 1$, then the G-R model predicts that there is only a 1 per cent chance that the event will grow to be larger than M 7.5 (Fig. 1b). An EEW system based on the G-R model will always conclude that there is a very low probability that such an earthquake will grow to be a 'Big One'; only 1 out of 100 earthquakes with a present magnitude of M_p 5.5 will eventually grow to be larger than M 7.5. Assuming higher b-values in the G-R model will increase the probabilities of earthquake growth (Iwata *et al.* 2005).

The CE model, in contrast, assumes that single faults tend to produce earthquakes of similar size. The magnitudes of these 'characteristic' earthquakes mainly depend on the dimension of the fault zones or fault segments (Schwartz & Coppersmith 1994). It is typically supposed that characteristic earthquakes are accompanied by foreshocks, aftershocks and low-level background seismicity along the considered fault or fault zone, which are distributed according to the G-R model. Youngs & Coppersmith (1985) and Convertito *et al.* (2006) give a PDF of the earthquake occurrence that is appropriate for this model with

$$\mathrm{prob}_{\rm CE}(M) = \begin{cases} \frac{\beta \mathrm{e}^{-\beta(M-M_{\rm min})}}{1-\mathrm{e}^{-\beta(M_{\rm max}-M_{\rm min}-\Delta M_2)}} \frac{1}{1+c} & \text{for } M_{\rm min} \le M < M_c \\ \frac{\beta \mathrm{e}^{-\beta(M_{\rm max}-M_{\rm min}-\Delta M_1-\Delta M_2)}}{1-\mathrm{e}^{-\beta(M-M_{\rm min}-\Delta M_2)}} \frac{1}{1+c} & \text{for } M_c \le M \le M_{\rm max} \end{cases}, \tag{5}$$

where the constant c is given by

$$c = \frac{\Delta M_2 \beta \mathrm{e}^{-(M_{\rm max}-M_{\rm min}-\Delta M_{1-}\Delta M_2)}}{1-\mathrm{e}^{-\beta(M_{\rm max}-M_{\rm min}-\Delta M_2)}}. \tag{6}$$

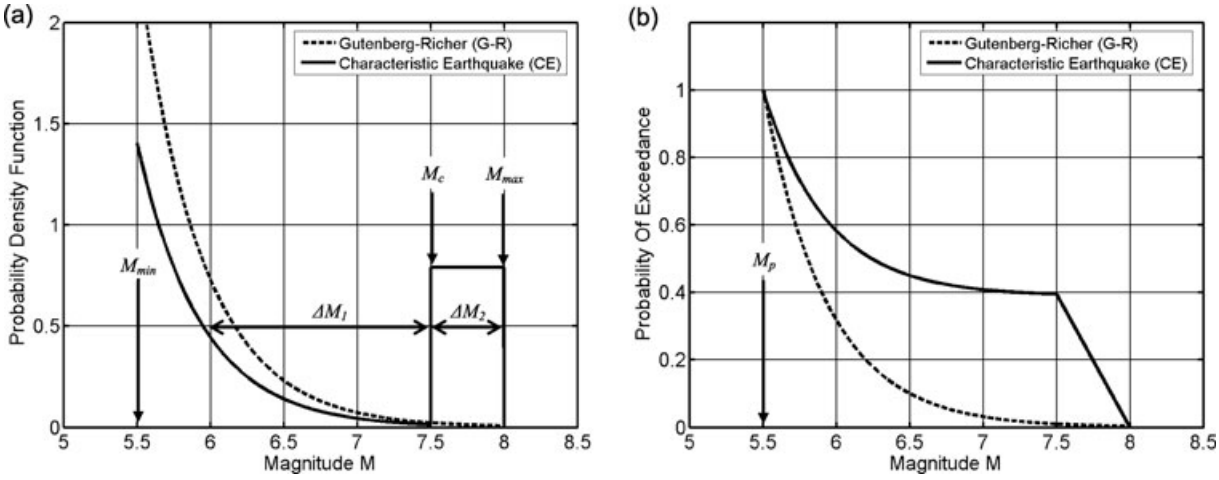

Figure 1. (a) Probability density function (PDF) of magnitude M and (b) probability of exceedance (POE) of different M for a rupture with a present magnitude of M_p 5.5. The probability for the occurrence of a large earthquake is much lower in the Gutenberg-and-Richter (G-R) than in the Characteristic Earthquake (CE) model: only 1 out of 100 earthquakes will exceed M 7.5. An EEW system based on the G-R model will likely never predict a 'Big One'. The CE model, on the other hand, is very poorly constrained (here: $b = 1.0$, $\Delta M_1 = 1.75$ and $\Delta M_2 = 0.5$).

© 2010 The Authors, *GJI*, **183,** 1014–1030
Geophysical Journal International © 2010 RAS

The parameters ΔM_1 and ΔM_2 describe two intervals below and above the characteristic magnitude $M_c = M_{\max}$–ΔM_2 (Fig. 1a).

The POE for the CE model, p_{CE}, can be determined analogous to eq. (4). Due to the occurrence of characteristic earthquakes, p_{CE} takes generally higher values than $p_{\mathrm{G\text{-}R}}$. For example with $\Delta M_1 = 1.75$ and $\Delta M_2 = 0.5$, the CE model predicts that there is a 40 per cent chance for an ongoing event with M_p 5.5 to grow to be larger than M 7.5, compared to 1 per cent predicted by the G-R model (Fig. 1b).

Spatially smooth ruptures along so-called 'mature' faults usually travel over larger distances than heterogeneous ruptures along 'immature' or 'generic' faults (e.g. Kanamori & Allen 1986; Wesnousky 1988; Manighetti *et al.* 2005, 2009). For this study we assume that the frequency-size statistics of earthquakes along mature faults can be described by the CE model, while the G-R model seems to be more appropriate to characterize the seismicity along generic faults. There is an ongoing debate among seismologists about which model might be more appropriate to characterize the seismic activity along individual faults (e.g. Wesnousky 1994; Kagan 1996; Page *et al.* in preparation). We want to avoid these discussions here. Even if we knew that the G-R or CE models are adequate to describe the seismicity in a certain region or along a specific fault, it seems clear that there is insufficient knowledge to assign appropriate parameters to the wide range of possible rupture scenarios.

A more flexible approach to study the rupture probabilities along faults is by the application of stochastic slip models that can be tuned from G-R- to CE-like behaviour. These models can be used to generate series of spatially variable slip (Liu-Zeng *et al.* 2005). Slip heterogeneities can be caused, for example, by the frictional properties and the spatial heterogeneities in the material characteristics on the fault, or by the complexities of the fault geometry (Rice 1993; Andrews 1994; Aki 1995; Aagaard & Heaton 2008). In this study, we use 1-D stochastic slip models to simulate large suites of possible rupture scenarios for a systematic statistical analysis of the relationship between seismic slip and rupture length.

We find that the larger the present slip amplitude D_p, the higher is the probability for the continuation of a rupture and the evolution into a 'Big One', that is, a long remaining rupture length L_r (Yamada 2007). We also find that the D_p–L_r relationship shows a significant amount of scatter and is strongly controlled by the degree of assumed spatial heterogeneity of slip on the fault.

In this paper, we are not addressing the origin of spatially heterogeneous slip. We simply note that if slip changes rapidly, then it is more difficult to predict the future evolution of the rupture (in space and time) than if the slip is spatially smooth. In principle, we could also try to estimate the smoothness of an ongoing rupture in real time from the variability of slip amplitudes relative to the mean evolution of slip. This idea will be briefly discussed in the next section, but shall not be a major topic in this study.

The real-time application of our probabilistic approach to EEW requires the continuous monitoring of seismic slip along faults, for example, by a global positioning system (GPS; e.g. Crowell *et al.* 2009) or by the backprojection of seismic displacement data onto the fault line (Yamada 2007). Again, to understand the implications of a small or large slip amplitude, it is essential to consider the respective 'smoothness' of the fault along which a detected earthquake propagates. This means that an EEW system for large earthquakes needs some mechanism for the rapid recognition of the rupturing fault and the capability to consider the respective statistical features. This procedure requires that the rupture is automatically linked to a (hopefully) known geological structure, for which a 'smoothness measure' from an earlier assessment was determined and stored in a database.

If rupture is occurring on a smooth fault, for example, on the San Andreas Fault (SAF), an EEW system could be implemented in two ways: the system could issue a warning, whenever slip on the fault exceeds a few metres, because the probability for a large earthquake is high and strong shaking is expected in large areas around the fault. A more sophisticated EEW system could use the present slip on the fault to estimate the final rupture dimensions and future slip evolution, and (using this information) could provide probabilistic predictions of seismic ground motions along the evolving rupture. We will demonstrate this concept of an EEW system on the example of the southern Californian SAF.

DATA

1-D stochastic slip models

We assume that seismic ruptures consist of slip pulses that propagate along a fault (Heaton 1990). The ultimate slip $D(x)$ at the spatial point x on the fault is achieved quickly after the passage of the rupture front. That is, we suppose that the current distribution of slip up to a present rupture length L_p is approximately the final (static) distribution of slip along this length, and no further slip accumulation occurs after the passage of the rupture front.

Although earthquakes actually occur on rupture surfaces, we are constructing here a simple simulation of long ruptures and assume a 1-D line source. We assume that slip pulses vary (perhaps chaotically) as they propagate along the fault. We also assume that the event is over, once the amplitude of the slip pulse drops to zero.

We simulate the propagation of slip pulses from simple 1-D stochastic slip models as proposed by Liu-Zeng *et al.* (2005), in which the slip amplitudes at adjacent points, x and $(x + \Delta x)$, change by some random amount in either positive or negative direction. We model this change in the amplitude by a low-pass filtered Gaussian white noise series $R(x)$ with zero mean and standard deviation one,

$$D(x) = \Delta D |FT^{-1}[\hat{R}(k)k^{-\mu}]|, \tag{7}$$

where $\hat{R}(k)$ is the Fourier transform of $R(x)$, k is the wavenumber, FT^{-1} the inverse Fourier transform and ΔD a scaling factor. The filtering exponent μ determines the spatial roughness of slip: the smaller μ, the rougher is the slip. Each point x, at which $D(x) = 0$, defines the end of an individual rupture. We take the absolute amplitudes of these 'parent series' to obtain contiguous positive slip functions (Fig. 2). Note that the slip models described by eq. (7) are purely geometric; they only assume that slip is spatially contiguous and that it has the scale-invariant properties of a self-affine fractal (Turcotte 1997).

Liu-Zeng *et al.* (2005) found that models with a rupture roughness of up to $\mu = 1.5$ produce slip series that have similar statistical features like real earthquakes (see below for further discussions). We use the stochastic slip models in eq. (7) to generate a set of 500 parent series of 65 536 points each for $\mu = 1.1$, $\mu = 1.3$ and $\mu = 1.5$. A total of 22 541 individual slip events, which are identified by determining adjacent zero crossings of the parent series, are catalogued: 15 006 for $\mu = 1.1$, 5121 for $\mu = 1.3$ and 2414 for $\mu = 1.5$.

Models with small rupture roughness μ, that is, high slip heterogeneity, tend to generate more small events, that is, shorter ruptures, than models with large μ (Fig. 3a), because rougher slip distributions are more likely to terminate rupture (Fig. 2). The frequency-size statistics of these events are close to the power-law distributions

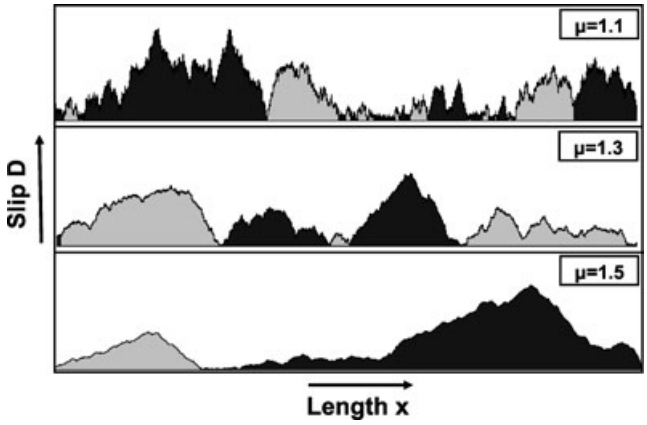

Figure 2. Unscaled parent series of slip generated from eq. (7) for three levels of rupture roughness μ. The larger the slip heterogeneity, the larger is the probability to cross the zero line, which defines the end of an individual rupture event.

generated by the G-R model; slip series generated from models with large μ, on the other hand, produce smooth ruptures with a higher number of large events, that is, the frequency-size statistics of these series are more similar to the distributions produced by the CE model (Fig. 3a). This means that we can use the roughness parameter μ to tune the stochastic models from G-R- to CE-like behaviour. This is in good agreement with, for example, Hillers *et al.* (2007), who found from quasi-dynamic rupture simulations that smooth continuum faults usually tend to have narrow event distributions, and that strong spatial heterogeneity is required to produce spatio-temporal non-uniform seismic behaviour.

The calibration of the 1-D slip functions is driven by two desires: (1) we want to reproduce realistic average slip-to-rupture length ratios as observed during earthquakes, and (2) we want to avoid maximum slip amplitudes that seem unrealistic high (although we are not really sure what is unrealistic). In this study, we set in eq. (7) $\Delta D = 3$ m. We assume that the 1-D slip amplitudes correspond to the average slip taken over the fault width.

The slip amplitudes D in our models follow one-sided Gaussian distributions (Fig. 3b) as a consequence of the assumed Gaussian distribution of noise amplitudes in eq. (7). If desired, we could replace the Gaussian by any other distribution, for example, by the logarithmic distribution. The slip amplitudes in the stochastic models increase with increasing rupture roughness, that is, decreasing μ (Fig. 3b). The largest slip amplitudes in our models are less than 30 m for $\mu = 1.3$ and $\mu = 1.5$, and, in a very few cases, grow up to 65 m for $\mu = 1.1$. The range of simulated slip amplitudes seems to be fairly realistic if compared, for example, with the slip models in the Finite-Source Rupture Model Database (SRCMOD version 7, http://www.seismo.ethz.ch/srcmod/). Note, however, that the inverted slip distributions in the SRCMOD are usually strongly smoothed and the SRCMOD does not contain a sufficient number of large events to allow for a statistical analysis as aimed for in our studies.

We generally find a good agreement between the ratios of the average slip $\overline{D}$ (taken over the entire rupture length L) to the rupture length L in the simulated slip functions and earthquakes which usually have $\overline{D}/L$ values within the wedge-shaped area in Fig. 4: the $\overline{D}/L$ ratios, which are related to the seismic stress drops, increase among small earthquakes compared to large events, if we consider the combined effect of different levels of rupture roughness (Fig. 4). Ruptures with rougher slip distributions (smaller μ) tend to produce higher slip-length ratios than smooth ruptures. Generally, the ratios decrease with increasing rupture length L, except for smooth ruptures ($\mu = 1.5$) that hardly show any variation in $\overline{D}/L$.

For rupture lengths that are large relative to the seismogenic width of the fault ($L \gg 30$ km), constant slip versus length scaling (of smooth ruptures) implies constant stress drops with L. This behaviour is compatible with the 'L-model' proposed by Scholz (1982). Note, however, that the stress drop is not really a fundamental frictional parameter in our 1-D stochastic slip models: the end of a rupture is simply defined by the location at which the amplitude of a slip pulse dies to zero. That is, the stress drop in our models is determined by the details of spatially variable slip pulses (Liu-Zeng *et al.* 2005).

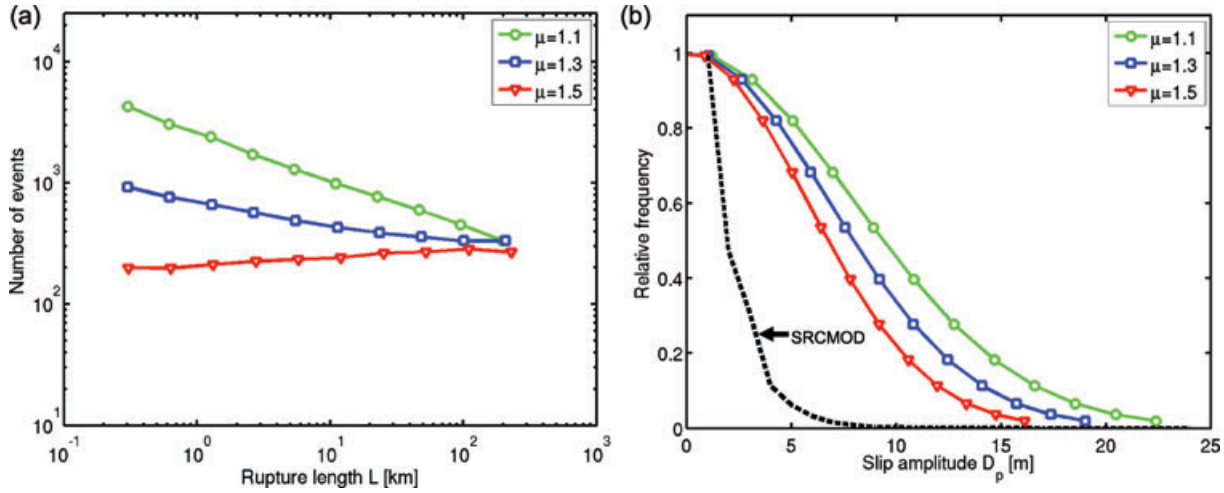

Figure 3. (a) Frequency-length and (b) frequency-slip statistics for the simulated slip events for three levels of rupture roughness μ using the 1-D stochastic slip models proposed by Liu-Zeng *et al.* (2005). Smooth ruptures (larger μ) tend to generate more long ruptures and smaller slip amplitudes than rough ruptures. The roughness parameter μ allows tuning the slip models from G-R- to CE-like behaviour. The range of simulated slip amplitudes agrees fairly well with the slip amplitudes in the Finite-Source Rupture Model Database (SRCMOD version 7; http://www.seismo.ethz.ch/srcmod/) (dashed line). Note, however, that the inverted slip distributions in the SRCMOD are usually strongly smoothed and the SRCMOD does not contain a sufficient number of large events to allow for a statistical analysis as anticipated in our study.

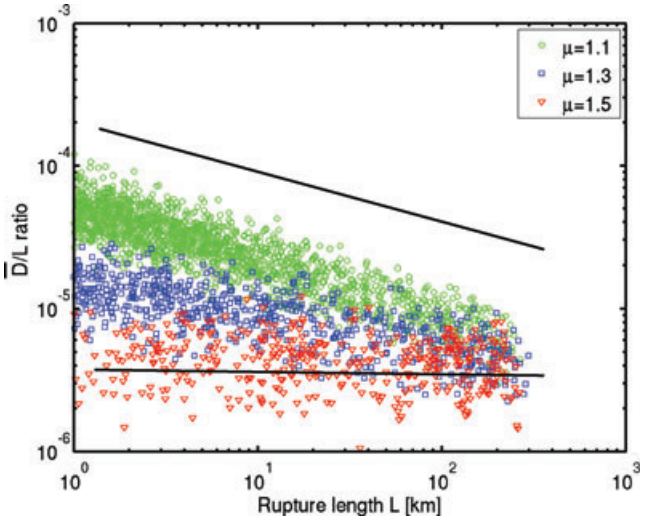

Figure 4. Average slip-to-rupture length $\overline{D}/L$ -ratios versus L obtained from the 1-D stochastic slip models for different levels of rupture roughness μ. The corresponding values of real earthquakes are typically located in the wedge-edged area between the two lines. The scaling of the parent series in eq. (7) is driven by the compromise between avoiding too large slip amplitudes and the desire to simulate ruptures with fairly realistic $\overline{D}/L$ ratios.

Generally, the $\overline{D}/L$ ratios obtained from the simulated slip series tend to be smaller than the ratios observed during real earthquakes (Fig. 4, wedge-edged area). In principle, we could improve the fit by increasing the scaling factor ΔD in eq. (7), but this would lead to single slip amplitudes exceeding several tens of metres, which seems too large for a real earthquake. The decision to set in this study $\Delta D = 3$ m is driven by the compromise between avoiding slip amplitudes that are too large and the desire to simulate ruptures with realistic $\overline{D}/L$ ratios.

Another way to evaluate the suitability of the simulated slip functions is by comparing their shape with observations made during real earthquakes. The mean shape of the 1-D slip functions can be characterized by three parameters (Ward 2004): the (total) rupture length L, the peak average slip E_{peak} and a skewness parameter q,

$$E(D_{L,E_{\text{peak}},q}(l)) = 2E_{\text{peak}}\frac{l^q}{L^{2q}}(L^{2q} - l^{2q})^{1/2}, \tag{8}$$

where q controls whether a slip function has more off-centre ($q \approx 2.0$) or more elliptical shape ($q \approx 0.5$; Ward 1997). We determine for each simulated slip series of length L parameters E_{peak} and q from fitting the model in eq. (8) to the generated 1-D slip functions (Figs 5a–d).

The majority of the slip series tend to have elliptical shapes, independent from the level of slip heterogeneity (Fig. 5c). This symmetry is not surprising since the probability for slip increase or decrease in the stochastic slip models is constant at all rupture lengths l, that is, on average we should obtain the same shape, independent from whether the rupture starts at $l = 0$ or at $l = L$. Manighetti *et al.* (2005) found that the slip functions of real earthquakes usually tend to be roughly triangular both along strike and dip, and most of them (70–80 per cent) are asymmetric. The reason for this phenomenon is not well understood, and cannot be reproduced by our simple static 1-D models.

We have seen that the 1-D stochastic slip models as used in this study imitate important statistical features of earthquakes, such as reproducing realistic slip amplitudes and average slip versus rupture length ratios (Fig. 4). On the other hand, the models fail to reproduce asymmetric shapes of the slip series as often are observed during real earthquakes. We think that employing the 1-D models for our purposes, however, is justified, because we are mainly interested in statistical features such as fractal behaviour of slip, which is well reproduced by the 1-D functions (see Discussions and Conclusions).

What is the relationship between the simulated slip series D in eq. (7) and the mean slip function E in eq. (8)? Neither the shape factor q nor the peak slip value E_{peak} show a clear dependency on μ (Figs 5c and d). However, the variability of the simulated slip amplitudes D relative to E increases strongly with increasing roughness of the slip series, that is, the stronger the slip heterogeneity, the larger are the differences of single slip amplitudes and the mean slip function, and the larger the standard deviation σ of the residuals (E-D) (Fig. 6).

From our stochastic 1-D models we find that $\sigma \leq 1.5$ m for $\mu = 1.5$, and $\sigma \leq 2.25$ m for $\mu = 1.1$ and $\mu = 1.3$ (Fig. 6). The variability increases with increasing mean slip E. We could in principle use the variability of the slip amplitudes along an ongoing rupture to estimate the rupture roughness in real-time. This procedure, however, requires densely sampled slip profiles. In the following, we assume that μ is a fault-characteristic parameter, which has been determined from previous geological and geophysical studies (e.g., Sagy *et al.* 2007) and is known when the rupture nucleates.

METHOD

Bayesian approach

The statistical analyses of seismic slip amplitudes D and rupture length L in this paper are based on the Bayesian theorem. Probabilistic methods, such as provided by the Bayesian framework, are reasonable approaches to deal with the large variability in the observed seismic source and ground motion characteristics and the inherent uncertainties in predicting these parameters.

In this study, we determine the probabilities of remaining rupture length L_r of an ongoing earthquake rupture using a presently observed seismic slip amplitude D_p. We are not only interested in determining the most probable, but also the less likely, however, still possible solutions. In other words: we want to determine the entire PDF for L_r. Both slip amplitudes D and rupture length L span several orders of magnitudes. They usually follow power-law distributions, and we therefore use their logarithmic values $\log(D)$ and $\log(L)$ in the following analyses.

We describe the rupture length–slip relationship by the 'conditional probability' of L_r for a given D_p, also known as the 'posterior probability' $p(\log(L_r)|\log(D_p))$. The Bayesian theorem states that the posterior probability can be determined from the product of the likelihood function $p(\log(D_p)|\log(L_r))$ and the *a priori* probability $p(\log(L_r))$, that is,

$$\begin{aligned} p(\log(L_r)|\log(D_p)) &= \frac{p(\log(D_p)|\log(L_r))p(\log(L_r))}{p(\log(D_p))} \\ &= \frac{\text{likelihood} \times a\ priori}{\text{normalization factor}}, \end{aligned} \tag{9}$$

where the normalization factor in the denominator, $p(\log(D_p)) = \sum_{i=1}^{n} p(\log(D_p)_i|\log(L_r))p(\log(L_r))$, ensures that the sum over all probabilities gives 1.

The likelihood function $p(\log(D_p)|\log(L_r))$ in eq. (9) describes, how likely it is to observe a certain $\log(D_p)$–$\log(L_r)$ combination.

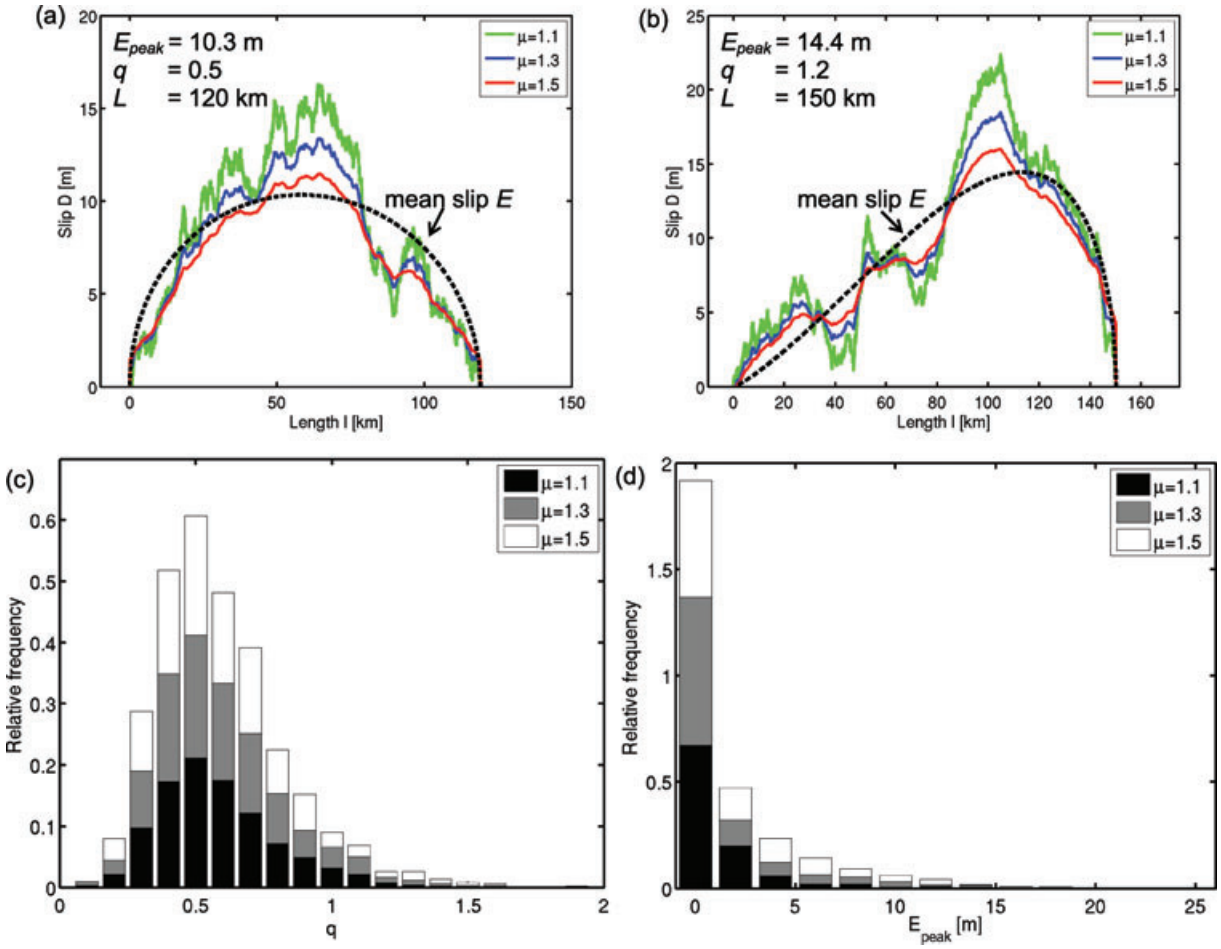

Figure 5. Statistical analyses of the shape of the simulated 1-D slip functions for three levels of rupture roughness μ. (a) and (b): Example slip series and mean slip functions E (dashed lines); E depends on the peak mean value E_{peak}, shape factor q and rupture length L (see eq. (8)). (c) and (d): Relative frequency of parameters q and E_{peak} for all 22 541 simulated ruptures. The majority of slip functions has a symmetric shape ($q = 0.5$) as a consequence of equal probabilities of slip increase and decrease at all length l, that is, in a statistical sense the slip series should not be different if the rupture starts at $l = 0$ or $l = L$.

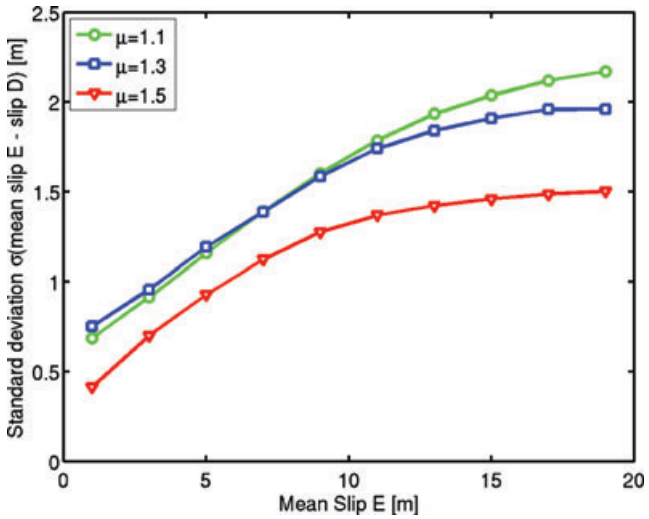

Figure 6. Standard deviation σ of residuals (mean slip E–slip D) for three levels of rupture roughness μ. The smaller the slip heterogeneity, the smaller is σ, that is, the better is the fit between E and D.

The *a priori* probability $p(\log(L_r))$ tells, how often we will observe a given $\log(L_r)$ value. Obviously, $p(\log(L_r))$ depends on both the probability of occurrence of a rupture with length L and on how far the considered rupture already has propagated, that is, $p(\log(L_r)) = p(\log(L–L_p))$, where L_p is the present rupture length, at which slip D_p is observed. In principle, we can use the log–linear relationship between the rupture length and magnitude in eq. (1) and describe the *a priori* probability $p(M)$ by the earlier discussed frequency-size statistics of earthquakes as produced, for example, by the G-R or the CE model.

The likelihood function $p(\log(D_p)|\log(L_r))$ in eq. (9) can be determined from the joint probability function $p(\log(D_p),\log(L_r))$, which describes the probability of the joint occurrence of $\log(D_p)$ and $\log(L_r)$. The likelihood function $p(\log(D_p)|\log(L_r))$ can be calculated from

$$p(\log(D_p)|\log(L_r)) = \frac{p(\log(D_p), \log(L_r))}{p(\log(L_r))}, \tag{10}$$

where we determine the joint probabilities from the 2-D kernel-based estimation method (e.g. Bishop 1995) using Gaussian kernels

with

$$p(\log(D_p), \log(L_r)) = \frac{1}{N}\frac{1}{\sigma\sqrt{2\pi}} \times \sum_{n=1}^{N} \exp\left(-\frac{\|(\log(D_p), \log(L_r)) - (\log(D_p), \log(L_r))_n\|^2}{2\sigma^2}\right), \quad (11)$$

where the standard deviation σ in the denominator controls the width of the kernels, and N is the number of data points. Kernel-based methods belong to the group of 'non-parametric' density estimation techniques that do not require assumptions about the forms of the PDFs (except that they are smooth). They therewith are highly flexible and well suited for many applications (e.g. Bishop 1995).

RESULTS

Probabilistic prediction of rupture length L and magnitude M

We go along each stochastically simulated 1-D slip series and determine for each slip amplitude D_p the PDF of remaining rupture length L_r, that is, the posterior probability $p(\log(L_r)|\log(D_p))$. We consider two levels of slip heterogeneity in the stochastic models: $\mu = 1.1$ (rough rupture) and $\mu = 1.5$ (smooth rupture) representing 'generic' and 'mature' fault conditions, respectively. Note that the slip and length values as specified in the following should be understood in a qualitative rather than a quantitative way, because they depend on the calibration of the 1-D slip models as described in the previous section.

Fig. 7 shows the probabilities of remaining rupture length L_r for slip amplitudes of $D_p = 2.0$ m, $D_p = 5.0$ m and continuous values of $D_p \leq 20.0$ m. The dashed lines mark the median values of the PDFs, that is, 50 per cent of the simulated ruptures propagate over a shorter and 50 per cent over a larger distance than specified by this value. Note the strong asymmetry of the PDFs, leading to significant differences between the median values and the maxima of the PDFs. The greyish colours mark the 38 per cent, 68 per cent and 95 per cent confidence intervals of the PDFs, corresponding to 0.5σ, 1σ and 2σ of a Gaussian distribution.

Our model estimates a 50 per cent chance that a rough rupture ($\mu = 1.1$) keeps growing for $L_r > 6$ km if $D_p = 2.0$ m, and for $L_r > 20$ km if $D_p = 5.0$ m (Fig. 7, left-hand side). Using eq. (1) these lengths correspond to magnitudes of $M > 5$ and $M > 6$, respectively. Note that a more accurate specification of M requires the knowledge of the present rupture length L_p. The probabilities for rupture continuation increase substantially, if the rupture travels along a smooth (mature) fault with little slip heterogeneity ($\mu = 1.5$; Fig. 7, right-hand side): 50 per cent of the ruptures keep growing for $L_r > 45$ km ($M \sim 7.0$) if $D_p = 2.0$ m, and for $L_r > 70$ km ($M > 7$) if $D_p = 5.0$ m. Our model thus predicts that the potential sizes of ruptures on heterogeneous (generic) and smooth (mature) faults may differ by more than one full magnitude unit for the same slip amplitude.

On a rough fault, L_r scales almost linearly with D_p (Fig. 7c, left-hand side), while the confidence intervals become broader, that is, the uncertainties in the prediction increase with increasing slip. Note that even for large slip amplitudes (e.g. $D_p = 6.0$ m) every second rupture does not exceed 25 km and therewith $M < 6.5$ (provided that D_p is observed at $L_p \approx 0$ km, i.e. close to the initiation point of the rupture); for 38 per cent of the earthquakes ruptures we observe $L_r < 50$ km, for 68 per cent $L_r < 75$ km and for 95 per cent $L_r < 125$ km. If the rupture, however, occurs along a smooth (mature) fault (Fig. 7c, right-hand side), we observe a strong increase of L_r for $D_p \leq 5.0$ m. The PDFs hardly change for larger D_p amplitudes, that is, the remaining rupture length is largely independent from the present slip and every second rupture will grow for $L_r > 65$ km, that is, $M > 7.0$.

Fig. 8 shows a comparison of the predicted probabilities of remaining rupture length L_r and the relative frequency of L_r determined from the slip models in the Finite-Source Rupture Model Database (SRCMOD version 7, http://www.seismo.ethz.ch/srcmod/) for slip amplitudes of $D_p = 1$–5 m. Note that these plots show observational slip data from both heterogeneous and smooth ruptures. There is qualitative agreement between these data and the probabilities derived from the stochastic models with slip heterogeneities of $\mu = 1.1$ and $\mu = 1.3$. The stochastic models capture essential features of the SRCMOD. A more detailed comparison, however, required a classification of the SRCMOD ruptures with respect to their levels of slip heterogeneity, which needs to be done in the future.

Prediction of mean slip function

Although the prediction of single slip amplitudes D along the evolving rupture may likely remain an elusive goal, we can try to estimate the mean slip function E using eq. (8). This involves some knowledge or the prediction of the total rupture length L, the peak mean slip E_{peak} and shape factor q. The estimation of these parameters requires some information on the evolution of slip. For example, if we knew the entire slip series up to rupture length L_p, we could estimate $L = L_p + L_r$ from eq. (9) and then determine E_{peak} and q from curve fitting of function E in eq. (8) to the observed slip amplitudes D. The shape of the slip function, however, remains elusive, if E_{peak} and q cannot be reasonably well resolved.

We test the robustness of the curve-fitting procedure to determine E_{peak} and q if the number of slip amplitudes is reduced by downsampling of data points in the simulated slip series. We use the slip amplitudes along the full rupture, that is, from $l = 0$ to $l = L$, and assume that the total rupture length L of each event is known. Fig. 9 shows the rms errors for both parameters as a function of data completeness C. A completeness of $C = 0.1$, for example, means that only every 10th slip amplitude of the original slip series is used.

We can easily recognize a rapid decrease of the rms errors in Fig. 9 with increasing data completeness C, that is, the more details (slip amplitudes) of the slip series are known, the more reliable are the estimates of q and E_{peak}. Rough ruptures (with small μ) that have a large variability in the slip amplitudes require more information on the evolution of slip than smooth series. Ruptures with slowly varying slip maintain their memory of past rupture over relatively longer lengths, that is, slip variations occur over larger length scales. We call this behaviour the (spatial) 'memory effect' of smooth ruptures.

In the previous example we assumed that the rupture is already terminated and slip amplitudes (or subsets of amplitudes) along this rupture as well as the total rupture length L are known, before q and E_{peak} are determined from curve fitting. In the case of an ongoing rupture, we will know the slip amplitudes only of the initial portion of the rupture (if at all). The results in Fig. 9, however, indicate that due to the (spatial) 'memory effect' of smooth ruptures the prediction of future slip does not require detailed information on

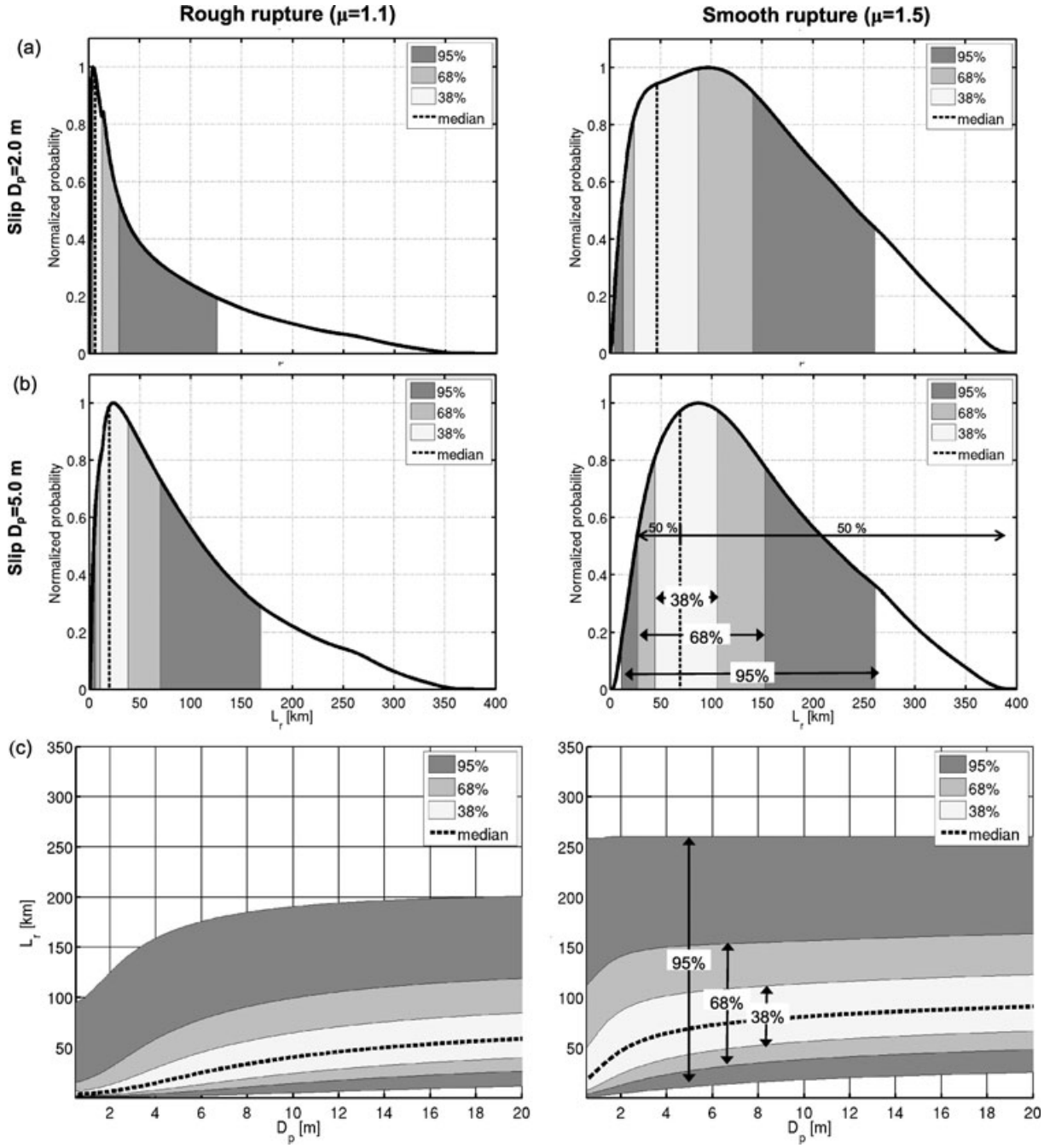

Figure 7. Probabilities of remaining rupture length L_r for a rupture with a present slip amplitude of (a) $D_p = 2.0$ m, (b) $D_p = 5.0$ m and (c) continuous values of $D_p \leq 20.0$ m. The probabilities are derived from the similated 1-D rupture series. Shown are the median values of the probability density functions (PDFs), as well as the 38 per cent, 68 per cent and 95 per cent confidence intervals, corresponding to 0.5σ, 1σ and 2σ of a Gaussian distribution. Note that the probabilities are plotted on a linear length scale, while determined on a logarithmic scale using eq. (9). Rough ruptures with a high level of slip heterogeneity (left-hand side) tend to propagate over shorter distances than smooth ruptures (right-hand side). The median values of the PDFs increase almost linearly with increasing slip D_p for a rough rupture (c, left-hand side), while staying almost constant after a rapid increase up to $D_p = 5.0$ m for a smooth rupture (c, right-hand side).

the evolution of slip if a smooth (mature) fault is considered. We will come back to this when demonstrating the probabilistic prediction for a scenario earthquake along the SAF.

Implications for EEW

EEW requires a rapid and robust determination of seismic source and ground motion parameters, such as magnitude and PGV, shortly after the initiation of a moderate-to-large earthquake. Most challenging in EEW for large earthquakes is the real-time determination of the probability that the rupture of an ongoing earthquake with a present length L_p (corresponding to a present magnitude M_p) will continue and exceed a certain magnitude level. We have seen earlier, that in the G-R model the POE for a large magnitude earthquake is extremely small: only 1 out of 100 earthquakes with M_p 5.5 ($L_p \approx$ 6 km) will exceed M 7.5 (Fig. 1). The output of the CE model, on the other hand, depends very much on the model parameters chosen and thus is not well constrained.

© 2010 The Authors, *GJI*, **183,** 1014–1030
Geophysical Journal International © 2010 RAS

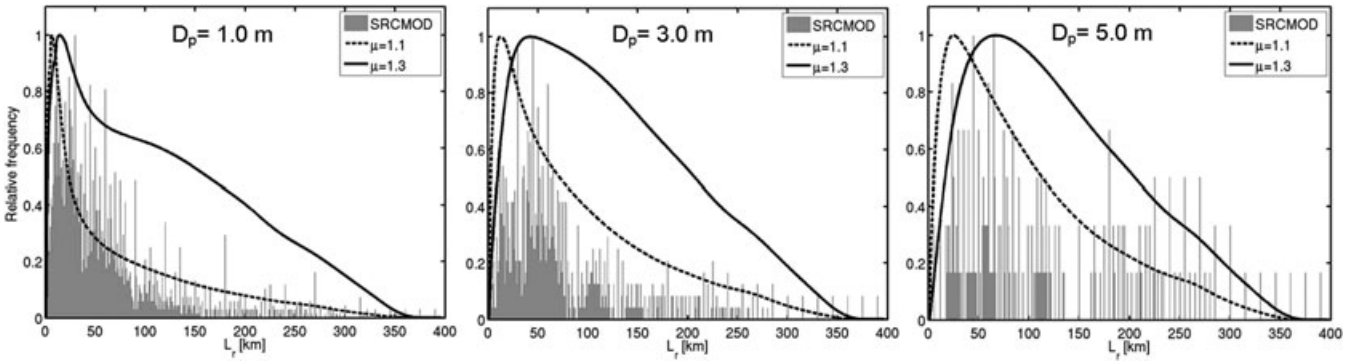

Figure 8. Comparison of the predicted probabilities of remaining rupture length L_r for two levels of slip heterogeneity μ with the relative frequency of L_r determined from the slip models in the Finite-Source Rupture Model Database (SRCMOD, http://www.seismo.ethz.ch/srcmod/); slip amplitudes vary from $D_p = 1$–5 m from left to right. Note that these plots show observational slip data from both heterogeneous and smooth ruptures. There is good agreement between these data and the probabilities derived from the stochastic slip models with $\mu = 1.1$ and 1.3. A more detailed comparison required a classification of the SRCMOD ruptures with respect to their levels of slip heterogeneity.

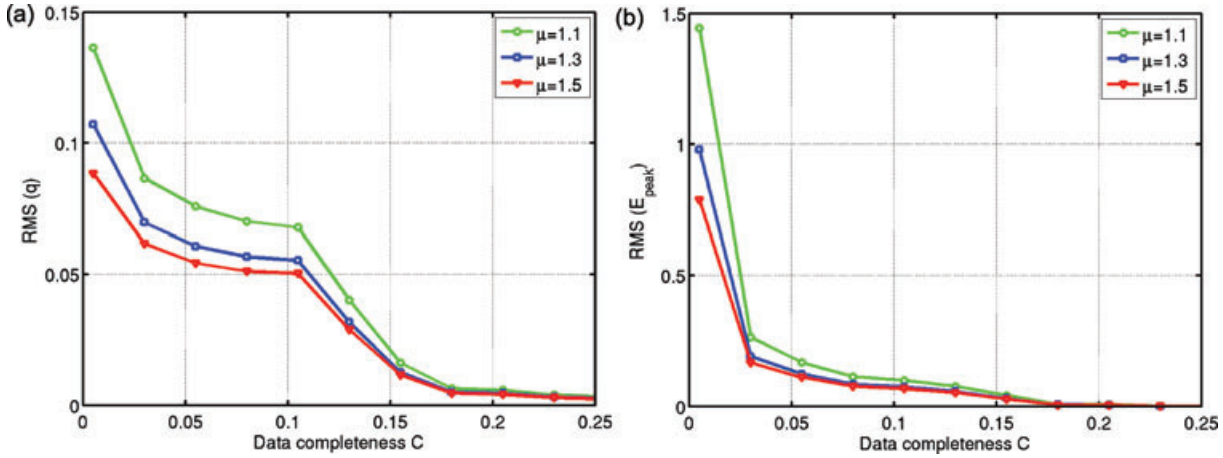

Figure 9. Robustness of the curve-fitting procedure to determine (a) the shape parameter q and (b) the peak average slip E_{peak} of the mean slip function E in eq. (8) if only a subset of slip amplitudes along the rupture is used. The rms errors decrease rapidly with increasing data completeness up to $C = 0.2$, which means that every fifth slip amplitude is assumed to be known. The larger the slip heterogeneity, the larger are the errors, that is, the more important is the knowledge of past slip for predicting future amplitudes. Ruptures with slowly varying slip (higher μ) maintain their memory of past rupture over relatively longer lengths. This is the so-called (spatial) 'memory effect' of smooth ruptures.

For the stochastic slip models the POE of a given magnitude tends to increase with increasing slip amplitude D_p (Fig. 10). For example, for M_p 5.5 and a slip amplitude of $D_p = 0.5$ m it is three times more likely that the rupture on a rough fault ($\mu = 1.1$) will become $M > 7.0$ and 10 times more likely to become $M > 7.5$ than in the G-R model (Table 1). The probabilities increase by an additional factor of two, if the slip increases to $D_p = 5.0$ m. A further increase of the probabilities by an additional factor of two to three is observed, if the same slip occurs along a smooth fault ($\mu = 1.5$). This observation demonstrates that although the present slip amplitudes D_p affect the POE, the smoothness of the causative fault has an even larger impact.

When should an EEW system issue a warning that an earthquake may become large? We have seen in Fig. 7c that L_r scales almost linearly with D_p on a rough rupture (left-hand side) and even for large slip amplitudes (e.g. $D_p = 6.0$ m) only every second rupture exceeds $L_r = 25$ km. If the rupture, however, occurs along a smooth (mature) fault, we observe a very strong increase of L_r with increasing $D_p \leq 5.0$ m (Fig. 7c, right-hand side). The PDFs hardly change for $D_p > 5.0$ m, suggesting that the remaining rupture length is largely independent from the present slip and every second rupture will grow $L_r > 65$ km, that is, $M > 7.0$.

These findings imply that we might have a critical slip amplitude (here $D_{p,\text{critical}} \sim 5.0$ m) on a smooth (mature) fault, for which we have an increased probability for the occurrence of a large earthquake for which EEW is needed. For the ruptures along heterogeneous (generic) faults, however, we cannot identify a critical slip value (Fig. 7c, left-hand side). Note again, that $D_{p,\text{critical}}$ as specified here depends on the scaling of the 1-D slip models, and might not be equal to the critical slip amplitude that we have to expect for a real earthquake. A more accurate quantification of $D_{p,\text{critical}}$ will require further statistical analyses of earthquake data and modelling of earthquake ruptures that is beyond the scope of this paper.

The statistical analyses of slip and rupture length in this paper demonstrate that a large slip amplitude D_p does not necessarily mean that we have a high probability for a large magnitude earthquake. Or, a small-to -moderate slip amplitude does not necessarily mean that the probability for a large event is low. We need to consider the respective characteristics of the underlying fault. We conclude from this finding, that an EEW system for large earthquakes requires a mechanism for the rapid recognition of the causative fault and consideration of its smoothness, that is, slip heterogeneity.

The majority of the algorithms, which currently are applied to EEW, use a fixed time window of a couple of seconds of the seismic

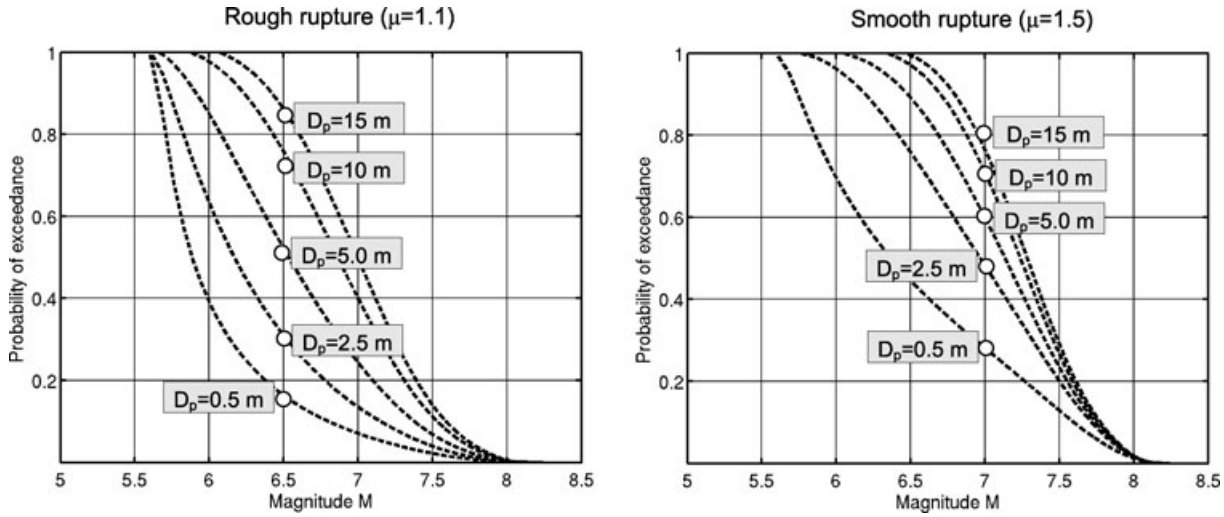

Figure 10. Probability of exceedance (POE) of different magnitude levels M for an ongoing rupture with a present magnitude of M_p 5.5 and present slip amplitude D_p. The probabilities are derived from the 1-D stochastic slip models for rough (left-hand side) and smooth ruptures (right-hand side), respectively. The POE is controlled by the slip amplitudes, D_p, the smoothness of the causative fault, however, has an even larger impact. The probabilities for a large event are much higher than in the G-R model (Fig. 1).

Table 1. Probability of exceedance (POE) of different magnitude levels M for a rupture with a present magnitude of M_p 5.5. The POE is derived from 1-D stochastic slip models with two levels of rupture roughness μ and slip amplitudes D_p. The probabilities are higher than predicted by the Gutenberg-and-Richter (G-R) model.

		Stochastic slip models with $\mu = 1.1/1.5$ (rough/smooth rupture) and slip D_p				
Magnitude	G-R model	$D_p = 0.5$ m	$D_p = 2.5$ m	$D_p = 5$ m	$D_p = 10$ m	$D_p = 15$ m
$M \geq 6.0$	0.3	0.4/0.7	0.6/1.0	0.9/1.0	1.0/1.0	1.0/1.0
$M \geq 6.5$	0.1	0.2/0.5	0.3/0.8	0.5/0.9	0.8/1.0	0.9/1.0
$M \geq 7.0$	0.03	0.1/0.3	0.1/0.5	0.2/0.6	0.4/0.7	0.5/0.8
$M \geq 7.5$	<0.01	<0.1/0.1	<0.1/0.2	0.1/0.2	0.1/0.3	0.1/0.3

P wave for a rapid estimation of the earthquake magnitude, using, for example, the average or the predominant period of shaking (e.g. Nakamura 1988; Allen & Kanamori 2003; Kanamori 2005; Lewis & Ben-Zion 2007; Wu *et al.* 2007; Lewis & Ben-Zion 2008; Böse *et al.* 2009a,b). This approach is heavily disputed, since the rupture process of earthquakes with $M > 6$ takes much longer than 3 s and it appears questionable, why the eventual rupture length and therewith the magnitude of an earthquake should be pre-determined at this early stage of the rupture process (e.g. Rydelek & Horiuchu 2006; Rydelek *et al.* 2007). Only if a large amount of the seismic energy of the earthquake is released at its initial stage, that is, only if seismic ruptures start in patches of large seismic slip (as proposed, e.g. by Mai *et al.* 2005 or Manighetti *et al*. 2005), a fast and robust magnitude prediction might become feasible. If this is not the case, the magnitude will likely remain underestimated, if we use a limited time window of a few seconds only.

Our analyses might shed some light onto these discussions: it seems appropriate to assign to a smooth (mature) fault, such as the SAF, a higher probability that a rupture will continue and finally evolve into a 'Big One', than to a heterogeneous (generic) fault. This implies that the observation of a (apparent) moderate earthquake (judging from the first few seconds of waveform data) on a smooth (mature) fault should result in a stronger warning.

Once the probabilities of the total rupture length L and the eventual magnitude M of an ongoing earthquake are estimated (provided that L_p or M_p, respectively, are known), we can predict in a probabilistic manner the level and distribution of ground shaking amplitudes, such as PGV, along the evolving rupture using empirical magnitude/distance relationships (see, e.g. *Next generation attenuation*—special issue by *Earthquake Spectra*, 2008). The consideration of rupture dimensions allows for much more realistic assessments of the expected ground shaking than could be delivered by a point source approximation, in which the ground motion amplitudes are assumed to decay with increasing epicentral rather than rupture-to-site distance.

Peak ground displacement (PGD) [or inferred from this the response spectral acceleration (SA) at long periods], on the other hand, shows a good correlation with the static slip on the fault (e.g. Aagaard *et al.* 2001). This means that if we were able to predict the future slip of the evolving rupture, we could use this information for a significant improvement of the estimates of parameters PGD and SA for EEW.

Demonstration for an M 7.8 scenario earthquake along the southern SAF

We are going to demonstrate the probabilistic prediction of seismic ground motion and future slip for an M 7.8 scenario earthquake along the southern SAF in California (Fig. 11a, left-hand side). The SAF is a right-lateral strike-slip fault with a total length of about 1200 km. The average time interval between significant ruptures along the SAF varies from about 25 yr in the Cholame Valley to

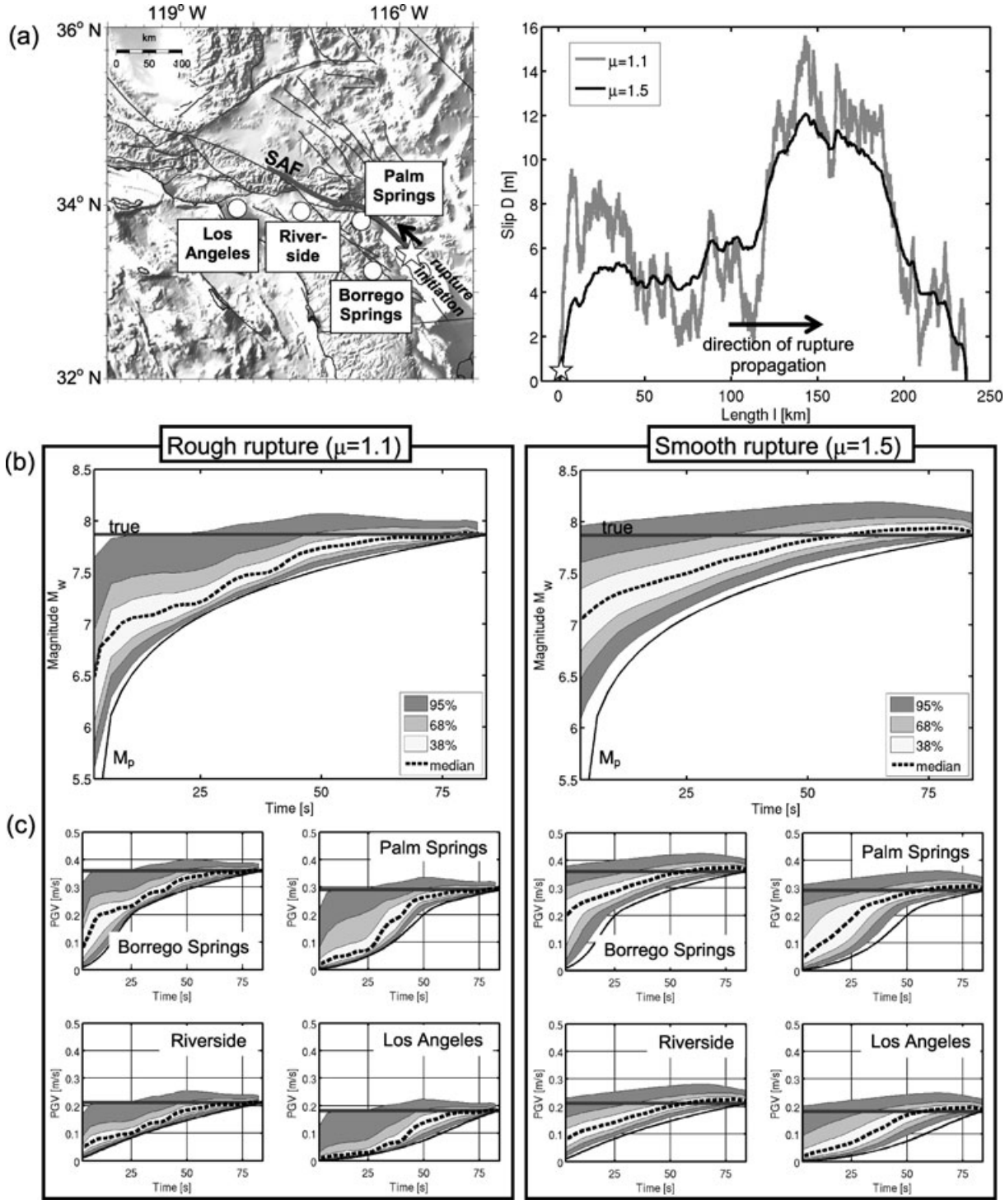

Figure 11. Scenario M 7.8 earthquake along the southern San Andreas Fault (SAF): (a) Rupture trace and distribution of four cities as potential recipients of EEW (left-hand side); simulated 1-D slip function for rough and smooth rupture conditions, respectively (right-hand side). (b) Probabilistic prediction of magnitude M as a function of time after rupture nucleation; shown are the 38 per cent, 68 per cent and 95 per cent confidence intervals as well as the median values of the PDFs; present magnitude M_p is determined from the present rupture length L_p; (c) Probabilistic prediction of peak ground velocity (PGV) at the four sites marked in (a). The convergence rate of the predicted values of M and PGV towards their true values is lower, if we assume a higher level of slip heterogeneity (left-hand side), that is, the severity of the event remains hidden for a long time of the rupture process.

about 300 yr on segments with very large slips. There is a 50 per cent chance of a magnitude M 7.5 or greater earthquake in California in the next 30 yr (www.scec.org/ucerf). The possible impact of an M 7.8 in southern California has been analysed and described by Jones *et al.* (2008). The warning time for a large earthquake on the SAF could be more than 1 min for large cities like Los Angeles that will experience a high level of shaking and damage during such an event.

Similar to the 'ShakeOut' scenario earthquake (Jones *et al.* 2008), the rupture of our scenario event starts on the southern portion of

the SAF, close to Bombay Beach, and propagates unidirectionally to the northwest over a distance of $L \sim 240$ km (Fig. 11a, left-hand side). We assume a constant rupture speed of $v_r = 2.8$ km s^{-1}, that is, the total rupture duration is ∼85 s. We simulate a 1-D seismic slip function of this scenario earthquake by usage of the earlier described stochastic slip models; that is, from our large collection of stochastic slip models, we choose several models with spatial distributions of slip that are qualitatively similar to the hypothesized 'ShakeOut' event. The SAF might be considered as a smooth fault with little slip heterogeneity. Thus, setting $\mu = 1.5$ in eq. (7) might be adequate. For comparison, however, we also model and analyse a scenario rupture of the same length L and similar slip D along a heterogeneous (generic) fault ($\mu = 1.1$). The maximum slip amplitudes are in the order of 15 m (Fig. 11a, right-hand side). By modelling purely unidirectional rupture propagation, we assume that the scenario earthquake starts with (almost) zero slip. The ruptures of real earthquakes often appear to propagate bidirectional and often start in patches of significant slip (e.g. Mai *et al.* 2005). EEW for unidirectional ruptures, as assumed in this scenario, is usually considered to be more difficult because of the small slip amplitudes at the beginning of the rupture. On the other hand, unidirectional rupture scenarios offer the largest possible warning times for EEW.

We use the earlier determined posterior probabilities $p(\log(L_r)|\log(D_p))$ in eq. (9) to resolve for each slip amplitude D_p in Fig. 11a (right-hand side) the PDFs of remaining and total rupture length, L_r and $L = L_p + L_r$, as a function of time after the rupture nucleation of the scenario earthquake. We then apply eq. (1) to estimate the corresponding magnitude M.

Obviously, the convergence rate of the estimated magnitude towards the true magnitude (M 7.8) is controlled by the slip heterogeneity on the causative fault: the median value of PDF is initially underestimated by 1.5 magnitude units, if slip occurs along a rough fault (Fig. 11b, left-hand side). Since on a heterogeneous (generic) fault, a short rupture is always more likely to occur than a large one, the predictions converge only slowly towards M 7.8 with increasing time and rupture growth. The median values are only slightly higher than the present magnitude M_p derived from the present rupture length L_p. If we assume, in contrast, that slip occurs along a smooth fault ($\mu = 1.5$), the median values of the magnitude PDFs are initially underestimated by less than one magnitude unit only (Fig. 11b, right-hand side), and show a fairly rapid convergence towards M 7.8.

In the next step, we are going to estimate the PGV in four cities along the SAF as potential recipients of EEWs: Borrego Springs, Palm Springs, Riverside and Los Angeles (Fig. 11a, left-hand side). At each time step after the rupture nucleation of the scenario earthquake, we determine PGV (for rock condition) from an empirical attenuation relation for earthquakes in southern California (Cua & Heaton 2008, submitted) using the predicted magnitudes in Fig. 11b and the respective distances of each user site to the present and predicted rupture along the SAF. A more detailed analysis could include the variability of PGV predictions due to the use of different attenuation relations into the Bayesian framework (see Discussions and Conclusions).

The scenario earthquake produces damaging shaking in large areas of southern California (Fig. 11c). At two out of four sites, Borrego Springs and Palm Springs, PGVs exceed 0.3 m s^{-1}. The other two sites, Riverside and Los Angeles, still experience PGVs of around 0.2 m s^{-1}. Similar to the magnitude M (Fig. 11b), PGV is initially underestimated for both rough and smooth rupture conditions (Fig. 11c). For the rough rupture (Fig. 11c, left-hand side), the predictions are once again only slightly better than if we base the prediction on the present rupture length L_p (rather than predicting $L = L_p + L_r$). The convergence is faster, if we assume that slip occurs along a smooth fault (Fig. 11c, right-hand side).

As said earlier, certain (mid- and long-period) ground motion parameters, such as PGD or SA, show a stronger correlation with the static slip on the fault than with earthquake magnitude. If the eventual length L of the rupture is known (or predicted), we can fit the mean slip function E in eq. (8) to the observed slip amplitudes D. We have seen in Fig. 9 that the required information on previous slip for predicting E decreases with increasing rupture smoothness due to the spatial 'memory effect' of smooth faults. In the following analyses we will assume the idealized case that the entire slip history up to the present rupture length L_p is known.

Fig. 12 shows the probabilistic prediction of the mean slip function E for the scenario M 7.8 earthquake at three time steps after the rupture nucleation (from top to bottom). The shape and peak parameters of the mean slip functions, q and E_{peak}, are determined from curve fitting of the 'known' slip amplitudes (in the grey area) to eq. (8) using the probabilistically predicted remaining rupture length L_r. The red line shows the mean slip function for the median value of the PDF of predicted L_r, the green and blue lines refer to the predicted rupture length within the 68 per cent and 95 per cent confidence intervals, respectively.

During the first two time steps (Fig. 12, figures in the top and middle panels), the peak mean slip E_{peak} is underestimated by almost a factor of two (5 m rather than 10 m) due to the fairly small slip amplitudes at the beginning of the rupture. The estimates of E improve, once the half of the rupture is finished. Note again that the predicted values of L_r are much smaller on the rough fault (Fig. 12, left-hand panels) than on the smooth fault (Fig. 12, right-hand panels), which has a strong impact on the shape parameter q: for the rough rupture it is $q < 0.5$, that is, E_{peak} is predicted to be close to $l = 0$. For the smooth (mature) fault, in contrast, we observe $q \geq 0.5$, that is, the slip function is assumed to be either symmetric or E_{peak} is predicted to be close to $l = L$ (elliptical to skewed shape). The smoother the rupture, the smaller are the differences between the predicted mean slip functions and the observed slip amplitudes.

Since we assume that the SAF can be characterized as a 'smooth' fault, the results shown on the right-hand side of Figs 11 and 12 seem to be more appropriate for this scenario. We conclude that a system for EEW for large earthquakes on the SAF (or on any other smooth fault with little slip heterogeneity) can be implemented in two ways: (1) The system can either provide real-time predictions of magnitude, slip and seismic ground motions with the corresponding uncertainties, whereby these estimates are steadily updated using observations of the present rupture length and present slip (Figs 11 and 12). Or, (2) The mere observation of slip in the order of a few metres in amplitude can trigger the declaration of a warning, since large slip on a smooth fault suggests a high probability of a large earthquake and strong shaking in wide areas around the fault (Fig. 7c). In contrast to the first implementation, the second EEW system will not provide detailed estimates of expected ground motions.

DISCUSSIONS AND CONCLUSIONS

An EEW system that estimates the potential source dimensions of a currently ongoing earthquake rupture using the present magnitude M_p together with the G-R frequency-size statistics will likely never predict a large magnitude earthquake, because of the rare occurrence of 'extreme events'. For an ongoing rupture with a present length

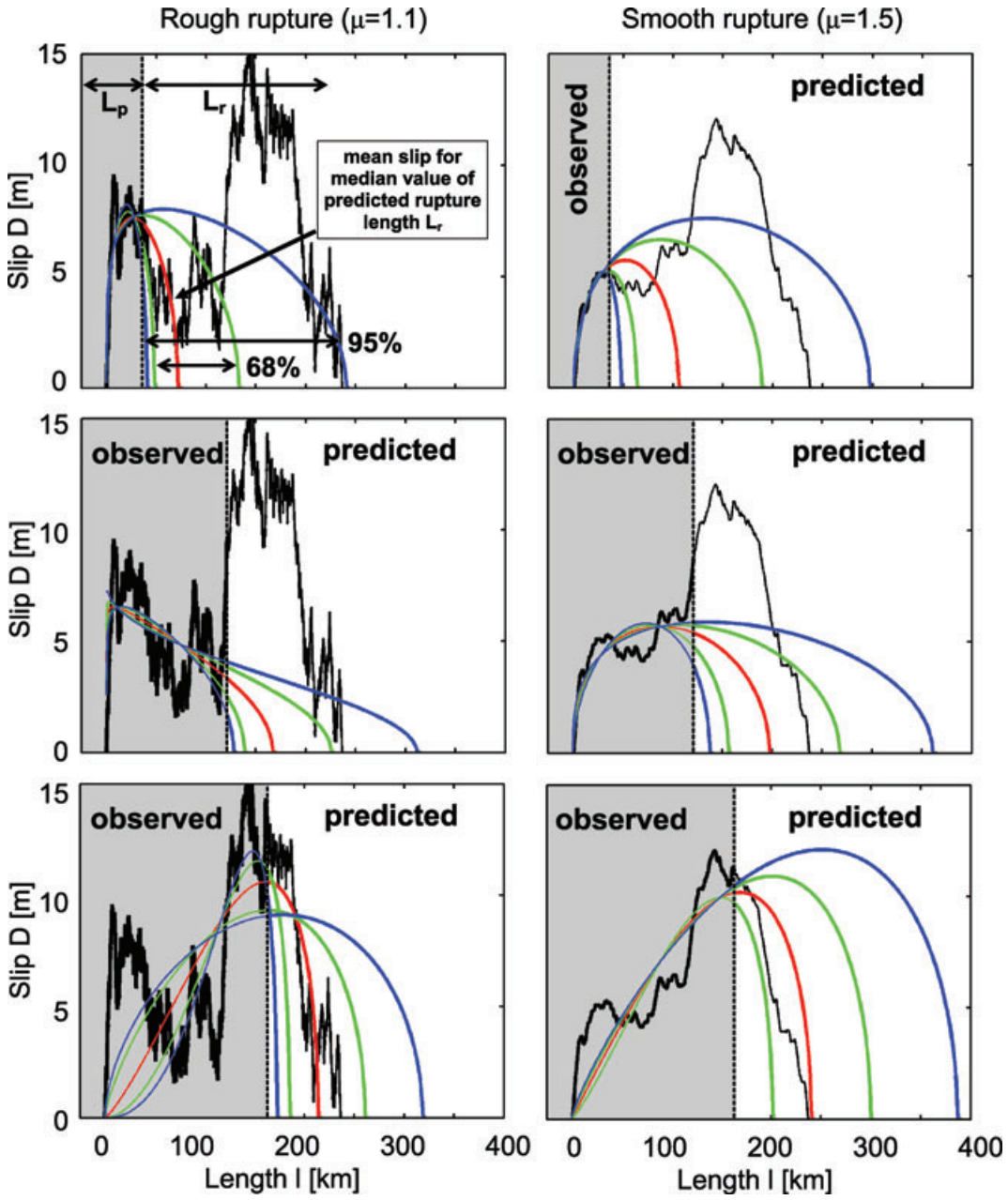

Figure 12. Probabilistic prediction of the mean slip function E for the scenario M 7.8 earthquake in Fig. 11 at three time steps after the rupture nucleation (from top to bottom), assuming a rough (left-hand side) and a smooth rupture (right-hand side) with high and low slip heterogeneity, respectively. The shape and peak parameters of the mean slip functions, q and E_{peak}, are determined from curve fitting of the 'known' slip amplitudes (in the grey area) to eq. (8) using the probabilistically predicted remaining rupture length L_r. The red line shows the mean slip function for the median values of the PDF of predicted L_r, the green and blue lines refer to the predicted rupture length within the 68 per cent and 95 per cent confidence intervals, respectively. The smoother the rupture, the smaller are the differences between E and the observed slip amplitudes.

corresponding to, for example, M_p 5.5, the probability to exceed M 7.5 is only 1 per cent. However, it seems to be reasonable to assume that the evolution into a 'Big One' is supported by large slip amplitudes and a high level of 'smoothness' on the causative fault.

To investigate the relationship between the slip and the eventual size of an ongoing rupture, we simulated in this study suites of evolving ruptures using 1-D stochastic models of spatially heterogeneous slip. We found that while large slip amplitudes D_p increase the probability for the continuation of a rupture and the possible evolution into a large earthquake, the recognition that rupture is occurring on a spatially smooth fault with low slip heterogeneity has an even stronger impact. The higher the variability in the slip amplitudes along the evolving rupture, the harder it is to predict whether the trend of declining slip amplitudes suggests that a rupture is likely to cease or whether this decrease might be followed by another increase in slip, caused, for example, by the rupture of a further asperity.

We conclude that an EEW system for large earthquakes needs some mechanism for the rapid recognition of the causative fault (e.g. from a GPS) and for consideration of its 'smoothness'. For earthquake ruptures occurring on a smooth fault, for example, on the SAF, an EEW system could be implemented in two ways: the

system could issue a warning, whenever slip on the fault exceeds a few metres, because the probability for a large earthquake is high (Table 1) and strong shaking is expected to occur in large areas around the fault. A more sophisticated EEW system could use the present slip on the fault to estimate the final rupture dimensions and future slip evolution, and (using this information) could provide probabilistic predictions of seismic ground motions along the evolving rupture. The decision on whether an EEW system should be realized in the first or in the second way (or in a combination of both) is clearly user-specific, and depends on (1) the vulnerability of the considered facility, and (2) the costs in the case of under- or overestimated ground shaking (Cua & Heaton 2007; Grasso *et al.* 2007; Böse *et al.* 2008).

The possible integration of our probabilistic approach into an EEW system for large earthquakes is shown in Fig. 13. A regional monitoring (GPS and/or seismic sensor network) and processing system provides real-time measurements of slip along the rupturing fault (e.g. Crowell *et al.* 2009). From a regional fault model, the system identifies the fault or fault segment along which the rupture likely propagates. Depending on the corresponding fault characteristics (dimensions, generic/mature fault, probability of multi-segment propagation, *etc*.) and on the current slip amplitude D_p, the processing system evaluates the probability of remaining rupture length L_r. Combined with the present rupture length L_p, a probabilistic distribution of the expected total length $L = (L_p + L_r)$ is computed, and from this, using, for example, eq. (1), a probabilistic distribution of the magnitude M. Length and magnitude estimates are combined with empirical attenuation relationships for a probabilistic prediction of the seismic ground motions along the evolving rupture. As time progresses, additional data will become available and allow for updating these estimates.

The probabilistic real-time prediction of rupture expansions can improve EEW for large earthquakes by more realistic estimates of the levels and distributions of ground shaking along the earthquake rupture compared to point source approximations. Uncertainties are expressed by the specification of PDFs. The real-time prediction of the final rupture dimensions, slip and seismic ground motions along the ongoing rupture becomes feasible for EEW, if the rupture duration exceeds the processing time of the warning system. If we assume a constant rupture velocity of $v_r = 2.8$ km s^{-1}, it takes around 11 s for a 30-km-long rupture ($\sim M$ 6.5) to cease, 22 s for a 60-km-long rupture ($\sim M$ 7.0), $\sim$50 s for a 140-km-long rupture ($\sim M$ 7.5) and around 100 s for a 300-km-long rupture ($\sim M$ 8.0). This implies that the probabilistic rupture and ground motion prediction as outlined in Fig. 13 might become feasible for EEW for earthquakes with $M > 6.5$, which usually are considered to be most destructive.

Our findings might also shed some light onto the discussions of whether the eventual size of an earthquake is pre-determined at the initial stage of the rupture process, that is, of whether seismic ruptures are 'deterministic' (e.g. Olson & Allen 2005; Rydelek & Horiuchu 2006; Rydelek *et al.* 2007). This is indirectly assumed by many EEW algorithms when deriving seismic information on magnitude from the initial portion of the seismic *P* wave only (Nakamura 1988; Allen & Kanamori 2003; Kanamori 2005; Böse *et al.* 2007; Wu *et al.* 2007). Our results imply that the observation of a (apparent) moderate earthquake (judging from the first few seconds of seismic waveform data, e.g. from the predominant or the average period of shaking) on a smooth (mature) fault should result in a stronger warning than if observed on a heterogeneous (generic) fault, because of a higher probability of future strong shaking at sites far from the epicentre.

Hillers *et al.* (2006) simulated evolving ruptures in quasi-dynamic continuum models of a 2-D strike-slip fault in an elastic solid with different levels of geometrical heterogeneity tuned by the spatial variations of the critical slip distance L_c. They found that the nucleation phases of small and large events differ in models with heterogeneous L_c distributions [see also Ellsworth & Beroza (1995) and Beroza & Ellsworth (1996)]. Even if we do not analyse the seismic nucleation phase here, it seems that our results suggest

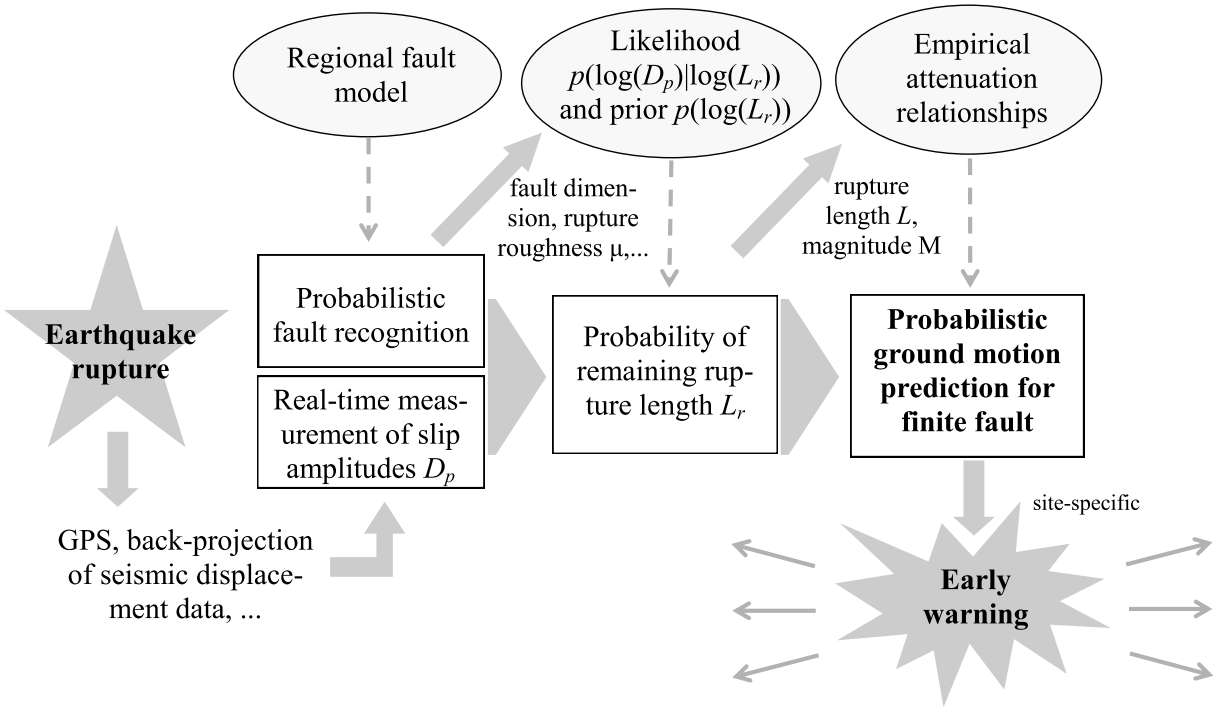

Figure 13. Possible integration of the probabilistic rupture prediction into a warning system for large earthquakes. Principle scheme.

that if ruptures occur on faults with very high heterogeneity, small and large events cannot be distinguished from each other at the early stage of the rupture. This observation is in good agreement with the model of cascading subevents on heterogeneous faults, in which the event size is influenced strongly by the highly variable stress-strength conditions along the fault (e.g. Ben-Zion *et al.* 2003).

The 1-D stochastic slip models used in this study were chosen for their conceptual simplicity, and there are indications that some of their features are incompatible with observations at real earthquakes. However, the 1-D models may capture several basic features of slip-pulse type models of dynamic rupture (Liu-Zeng *et al.* 2005). Other, more complex (dynamic) models might be better suited to reproduce time-dependent features of earthquakes, but they require the specification of a large number of free parameters (such as the pre-stress conditions on the fault) that are largely unknown. We therefore think that the usage of our static 1-D models in the presented study is justified by the fact that we describe the general features of slip, such as the dependency of rupture propagation on slip heterogeneity, in a statistical sense, rather than providing a quantitative analysis.

The root physical cause of heterogeneous slip is not well understood. It seems likely that geometric complexities in fault geometry and in material properties play a role (Rice 1993; Andrews 1994; Aki 1995). However, dynamic chaos that occurs when sliding friction includes strong slip velocity weakening may also be a source of slip heterogeneity (Aagaard & Heaton 2008). Regardless of the cause of the heterogeneity, our findings point to the desirability of characterizing spatial heterogeneity on recognized faults.

Manighetti *et al.* (2007) suggest that the variability in the slip-to-rupture length ratios arises from earthquakes breaking a variable number of fault segments with variable frictional strength that depends on both their long-term slip history ('structural maternity') and on their geometry (large-scale segmentation). This interpretation could provide a possible link between the concepts of slip roughness μ and fault maturity.

The classification of faults with respect to their level of maturity is not trivial. We hypothesize that a mature fault like the SAF could be identified from its seismicity (e.g. very few small earthquakes), its geological persistence, the total offset on the fault and its geometric simplicity (e.g. Aviles *et al.* 1987). Clearly, the number of mature faults worldwide is small, and we therefore suggest installing GPS monitoring systems along these few faults to build very simple and effective EEW systems.

In this study, we based the probabilistic prediction of L_r on the present slip amplitude D_p. We might possibly expect that also the average slip $\overline{D}$ taken between the spatial start point and the present rupture length $L_p = L - L_r$ might be suited for the prediction of L_r of an ongoing rupture. We found, however, that the correlation is smaller for $\overline{D}[R(\overline{D}, L_r) \approx 0.6]$ than for $D_p[R(D_p, L_r) \approx 0.8]$, implying that D_p is more suited for the prediction of L_r than $\overline{D}$. If the slip is very heterogeneous, then the rupture quickly loses its memory of past slip as it propagates along the fault. Ruptures with slowly varying slip maintain their memory of past rupture over relatively longer lengths (spatial 'memory effect' of smooth faults).

Even if we demonstrate our concept in this paper for a scenario earthquake with unidirectional rupture propagation, our probabilistic method is directly applicable to earthquakes with both uni- and bidirectional rupture propagation. Because of the symmetry of the problem, the variability in $p(\log(L_r)|\log(D_p))$ does not increase in the bidirectional case: assuming that the underlying fault is infinite, the estimates of remaining rupture length L_r to both directions are independent from each other and can be determined in the same manner as in the unidirectional case. However, the convergence to the true magnitude will be much faster in the bidirectional case and the distribution of ground motions might likely differ significantly compared to the unidirectional case. Bidirectional ruptures clearly complicate the problem of ground motion prediction and remain a field of future research.

Our probabilistic approach allows specifying inherent uncertainties in the estimated source and ground motion parameters. A complete assessment would require the additional consideration of, for example, the uncertainties in the rupture length–magnitude relation in eq. (1), uncertainties in the empirical attenuation laws, site and directivity effects. Yet another important issue in the context of rupture prediction clearly is the consideration of rupture propagation along different fault segments (possibly with different rupture roughness), as studied, for example, by Jackson *et al.* (1995) and Field *et al.* (2007). Including these effects in the probabilistic models as presented here, will lead to a further improvement of EEW for large magnitude earthquakes in the future.

ACKNOWLEDGMENTS

This work is funded through contract G09AC00258 from USGS/ANSS to the California Institute of Technology (Caltech). This is contribution #10039 of the Seismological Laboratory, Geological and Planetary Sciences at Caltech. Calculations and most figures in this paper were generated from Matlab 7.8. The map in Fig. 11a (left-hand side) was made using the Generic Mapping Tools version 4.2.1 (www.soest.hawaii.edu/gmt; Wessel & Smith 1998). We would like to thank Egill Hauksson for proofreading. We are grateful for the helpful comments of P. Martin Mai, two anonymous reviewers and the editor Yehuda Ben-Zion that helped us to improve an earlier version of this manuscript.

REFERENCES

Aagaard, B., Hall, J. & Heaton, T., 2001. Characterization of near-source ground motions with earthquake simulations, *Earthq. Spectra,* **17,** 177–207.

Aagaard, B. & Heaton, T., 2008. Constraining fault constitutive behavior with slip heterogeneity, *J. geophys. Res.,* **113,** B04301, doi:10.1029/2006JB004793.

Aki, K., 1995. Interrelation between fault zone structures and earthquake processes, *Pageop,* **145,** 647–676.

Allen, R.M., 2006 Probabilistic warning times for earthquake ground shaking in the San Francisco Bay Area, *Seismol. Res. Lett.,* **77**(3), 371–376.

Allen, R.M., Gasparini, P., Kamigaichi, O. & Böse, M., 2009. The status of earthquake early warning around the world: an introductory overview, *Seismol. Res. Lett.,* **80**(5), 682–693, doi:10.1785/gssrl.80.5.682.

Allen, R.M. & Kanamori, H., 2003. The potential for earthquake early warning in Southern California, *Science,* **300,** 786–789.

Andrews, D.J., 1994. Fault geometry and earthquake mechanics, *Annali Di Geofisica,* **XXXVII,** 1341–1348.

Aviles, C.A., Scholz, C.H. & Boatwright, J., 1987. Fractal analysis applied to characteristic segments of the San Andreas Fault, *J. geophys. Res.,* **92**(B1), 331–344.

Ben-Zion, Y., Eneva, M. & Liu, Y., 2003. Large earthquake cycles and intermittent criticality on heterogenous faults due to evolving stress and seismicity, *J. geophys. Res.,* **108**(B6), 2307, doi:10.1029/2002JB002121.

Beroza, G.C. & Ellsworth, W.L., 1996. Properties of the seismic nucleation phase, *Tectonophysics,* **261**(1–3), 209–227.

Bishop, C., 1995. *Neural Networks for Pattern Recognition,* Clarendon Press, Oxford.

Böse, M., 2006. Earthquake early warning for Istanbul using artificial neural networks, *Ph.D. thesis*, 181 pp., Karlsruhe University, Germany.

Böse, M., Hauksson, E., Solanki, K., Kanamori, H. & Heaton T. H., 2009b. Real-time testing of the on-site warning algorithm in southern California and its performance during the July 29 2008 M_w 5.4 Chino Hills earthquake, *Geophys. Res. Lett.,* **36,** L00B03, doi:10.1029/2008GL036366.

Böse, M., Hauksson, E., Solanki, K., Kanamori, H., Wu, Y.-M. & Heaton, T.H., 2009a. A new trigger criterion for improved real-time performance of on-site early warning in southern California, *Bull. seism. Soc. Am.,* **99**(2A), 897–905; doi:10.1785/0120080034.

Böse, M., Ionescu, C. & Wenzel, F., 2007. Earthquake early warning for Bucharest, Romania: novel and revised scaling relations, *Geophys. Res. Lett.,* **34,** L07302, doi:10.1029/2007GL029396.

Böse, M., Wenzel, F. & Erdik, M., 2008. PreSEIS: a neural network based approach to earthquake early warning for finite faults, *Bull. seism. Soc. Am.,* **98**(1), 366–382, doi:10.1785/0120070002.

Convertito, V., Emolo, A. & Zollo, A., 2006. Seismic-hazard assessment for a characteristic earthquake scenario: an integrated probabilistic-deterministic method, *Bull. seism. Soc. Am.,* **96**(2), 377–391.

Crowell, B.W., Bock, Y. & Squibb, M.B., 2009. Demonstration of Earthquake Early Warning using total displacement waveforms from real-time GPS networks, *Seismol. Res. Lett.,* **80**(5), 772–782, doi:10.1785/gssrl.80.5.772.

Cua, G. & Heaton, T., 2007. The Virtual Seismologist (VS) method: a Bayesian approach to earthquake early warning, in '*Earthquake Early Warning Systems*', pp. 85–132, eds Gasparini, P., Manfredi, G. & Zschau, J., Springer, Berlin, ISBN-13 978–3-540–72240-3.

Cua, G. & Heaton, T., 2008. Characterizing average properties of southern California ground motion amplitudes and envelopes, *Bull. seism. Soc. Am.,* submitted.

Ellsworth, W.L. & Beroza, G.C., 1995. Seismic evidence for an earthquake nucleation phase, *Science,* **268**(5212), 851–855, doi:10.1126/science.268.5212.851.

Field, E.H. & the Working Group on California Earthquake Probabilities, 2007. The uniform California earthquake rupture forecast, version 2, http://pubs.usgs.gov/of/2007/1437/ (last accessed 2010 February 5).

Goltz, J., 2002. Introducing earthquake early warning in California: a summary of social science and public policy issues, A report to OES and the Operational Areas, Caltech Seismological Laboratory, Disaster Assistance Division.

Grasso, V.F., Beck, J.L. & Manfredi, G., 2007. Automated decision procedure for earthquake early warning, *Eng. Struct.,* **29,** 3455–3463.

Gutenberg, R. & Richter, C.F., 1944. Frequency of earthquakes in California, *Bull. seism. Soc. Am.,* **34**, 185–188.

Heaton, T., 1985. A model for a seismic computerized alert network, *Science,* **228**(4702), 987–990.

Heaton, T.H., 1990. Evidence for and implications of self-healing pulses of slip in earthquake rupture, *Phys. Earth planet Int.,* **64,** 1–20.

Hillers, G., Ben-Zion, Y. & Mai, P.M., 2006. Seismicity on a fault controlled by rate- and state-dependent friction with spatial variations of the critical slip distance, *J. geophys. Res.,* **111,** B01403, doi: 10.1029/2005JB003859.

Hillers, G., Mai, P.M., Ben-Zion, Y. & Ampuero, J.-P., 2007. Statistical properties of seismicity of fault zones at different evolutionary stages, *Geophys. J. Int.,* **169,** 515–533.

Hoshiba, M., Kamigaichi, O., Saito, M., Tsukada, S. & Hamada, N., 2008. Earthquake early warning starts nationwide in Japan, *EOS, Trans. Am. geophys. Un.,* **89**(8), 73–74.

Iwata, T., Imoto, M. & Horiuchi, S., 2005. Probabilistic estimation of earthquake to a catastrophic one, *Geophys. Res. Lett.,* **32,** L19307, doi:10.1029/2005GL023928.

Jackson, D. & the Working Group on California Earthquake Probabilities, 1995. Seismic hazards in southern California: probable earthquakes, 1994 to 2024, *Bull. seism. Soc. Am.,* **85**(2), 379–439.

Jones, L.M. *et al*., 2008. The ShakeOut Scenario, U.S. Geological Survey Open File Report 2008–1150, California Geological Survey.

Kagan, Y.Y., 1996. Comment on "The Gutenberg-Richter or characteristic earthquake distribution, which is it?" by S. G. Wesnousky, *Bull. seism. Soc. Am.,* **86,** 274–285.

Kanamori, H., 2005. Real-time seismology and earthquake damage mitigation, *Annu. Rev. Earth planet. Sci.,* **33** 195–214, doi: 10.1146/annurev.earth.33.092203.122626.

Kanamori, H. & Allen, C.R., 1986. Earthquake repeat time and average stress drop, in *Earthquake Source Mechanics*, pp. 227–235, eds. Das, S., Boatwright, J. & Scholz, C., American Geophysical Union Geophysical Monograph, **37**.

Lewis, M.A. & Ben-Zion, Y., 2007. Examination of scaling between proposed early signals in P waveforms and earthquake magnitudes, *Geophys. J. Int.,* **171,** 1258–1268, doi: 10.1111/j.1365-246X.2007.03591.x.

Lewis, M.A. & Ben-Zion, Y., 2008. Examination of scaling between earthquake magnitude and proposed early signals in P waveforms from very near source stations in a South African gold mine, *J. geophys. Res.,* **113,** B09305, doi: 10.1029/2007JB005506.

Liu-Zeng, J., Heaton, T. & DiCaprio, C., 2005. The effect of slip variability on earthquake slip-length scaling, *Geophys. J. Int.,* **162,** 841–849, doi: 10.1111/j.1365-246X.2005.02679.x

Mai, P.M., Spudich, P. & Boatwright, J., 2005. Hypocenter locations in finite-source rupture models, *Bull. seism. Soc. Am.,* **95,** 965–980.

Manighetti, I., Campillo, M., Bouley, S. & Cotton, F., 2007. Earthquake scaling, fault segmentation, and structural maturity, *Earth planet. Lett.,* **253,** 429–438.

Manighetti, I., Campillo, M., Sammis, C., Mai, P.M. & King, G., 2005. Evidence for self-similar, triangular slip distributions on earthquakes: implications for earthquake and fault mechanics, *J. geophys. Res.,* **110,** B05302, doi: 10.1029/2004JB003174.

Manighetti, I., Zigone, D., Campillo, M. & Cotton, F., 2009. Self-similarity of the largest-scale segmentation of the faults: implications for earthquake behavior, *Earth planet. Sci. Lett.,* **288,** 370–381, doi: 10.1016/j.epsl.2009.09.040.

Nakamura, Y., 1988. On the urgent earthquake detection and alarm system (UrEDAS), in *Proceeding of the 9th World Conference on Earthquake Engineering* 1988, Tokyo-Kyoto, Japan.

Olson, E.L. & Allen, R.M., 2005. The deterministic nature of earthquake rupture, *Nature Lett.,* **438,** 212–215, doi: 10.1038/nature04214.

Rice, J.R., 1993. Spatio-temporal complexity of slip on a fault, *J. geophys. Res.,* **98**(B6), 9885–9907.

Rydelek, P. & Horiuchi, S., 2006. Is the earthquake rupture deterministic? Brief communications arising from: E.L. Olson and R.M. Allen, *Nature,* **44,** doi: 10.1038/nature04963.

Rydelek, P., Wu, C. & Horiuchi, S., 2007. Comment: peak ground displacement and earthquake magnitude, *Geophys. Res. Lett.,* **34,** L20302, doi: 10.1029/2007GL029387.

Sagy, A., Brodsky, E.E. & Axen, G.J., 2007. Evolution of fault-surface roughness with slip, *Geology,* **35**(3), 283–286, doi: 10.1130/G23235A.1.

Scholz, C.H., 1982. Scaling laws for large earthquakes: consequences for physical models, *Bull. seism. Soc. Am.,* **72,** 1–14.

Schorlemmer, D., Wiemer, S. & Wyss, M., 2004. Earthquake statistics at Parkfield I: stationarity of *b*-values, *J. geophys. Res.,* **109,** B12307, doi:10.1029/2004JB003234.

Schwartz, D.P. & Coppersmith, K.J., 1984. Fault behavior and characteristic earthquakes: examples from the Wasatch and San Andreas fault zones, *J. geophys. Res.,* **89,** 5681–5698.

Turcotte, D.L., 1997. *Fractals and Chaos in Geology and Geophysics,* Cambridge Univ. Press, New York.

Ward, S.N., 1997. Dogtails versus rainbows: synthetic earthquake rupture models as an aid in interpreting geological data, *Bull. seism. Soc. Am.,* **87,** 1422–1441.

Ward, S.N., 2004. Earthquake simulation by restricted random walks, *Bull. seism. Soc. Am.,* **94**(6), 2079–2089.

Wells, D.L. & Coppersmith, K.J., 1994. New empirical relationships among magnitude, rupture length, rupture width, rupture area and surface displacement, *Bull. seism. Soc. Am.,* **84,** 974–1002.

Wesnousky, S.G., 1988. Seismological and structural evolution of strike-slip faults, *Nature,* **335,** 340–343.

Wesnousky, S.G., 1994. The Gutenberg-Richter distribution or characteristic earthquake distribution: which is it? *Bull. seism. Soc. Am.,* **84,** 1940–1959.

Wessel, P. & Smith, W.H.F., 1998. New, improved version of the generic mapping tools released, *EOS, Trans. Am. geophys. Un.,* **79,** 579.

Wiemer, S. & Wyss, M., 2002. Mapping spatial variability of the frequency-magnitude distribution of earthquakes, *Adv. Geophys.,* **45,** 259–302.

Working Group on California Earthquake Probabilities, 1999. Earthquake probabilities in the San Francisco Bay Region: 2000 to 2030—A Summary of Findings, Open-File Report 99–517, US Geological Survey, http://geology.wr.usgs.gov/open-file/of99–517/of99–517.pdf.

Wu, Y.-M., Kanamori, H. Allen, R.M. & Hauksson, E., 2007. Determination of earthquake early warning parameters, τ_c and P_d, for southern California, *Geophys. J. Int.,* **170,** 711–717, doi: 10.1111/j.1365-246X.2007.03430.x.

Yamada, M., 2007. Early warning for earthquakes with large rupture dimension. *PhD thesis*, California Institute of Technology, Pasadena, 197pp., http://resolver.caltech.edu/CaltechETD:etd-03152007-174624

Yamada, M. & Heaton, T., 2008. Real-time estimation of fault rupture extent using envelopes of acceleration, *Bull. seism. Soc. Am.,* **98**(2), 607–619, doi: 10.1785/0120060218.

Yamada, M., Heaton, T. & Beck, J., 2007. Real-time estimation of fault rupture extent using near-source versus far-source classification, *Bull. seism. Soc. Am.,* **97**(6), 1890–1910, doi: 10.1785/0120060243.

Youngs, R.R. & Coppersmith, K.J., 1985. Implications of fault slip rates and earthquake recurrence models to probabilistic seismic hazard estimates, *Bull. seism. Soc. Am.,* **75,** 939–964.